

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 19:53:11 ; Search time 133 Seconds
(without alignments)
254.377 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCKAFQKSNLRHRT.....KQCKAFQCPNLRHGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	378	7 ADL33662	Adl33662 Zinc fing
2	340	76.9	107	9 AEB77956	Aeb77956 Zinc fing
3	340	76.9	111	8 ADP70949	Adp70949 Novel zin
4	334	75.6	111	8 ADP70959	Adp70959 Novel zin
5	330	74.7	111	8 ADP70948	Adp70948 Novel zin
6	328	74.2	111	8 ADP70930	Adp70930 Novel zin
7	328	74.2	111	8 ADP70950	Adp70950 Novel zin
8	328	74.2	113	8 ADP70936	Adp70936 Novel zin
9	328	74.2	125	8 ADP70936	Adp70936 Novel zin
10	327	74.0	213	5 AEO48443	Ado48443 Human VEG
11	327	74.0	483	5 AEG59963	Abg59963 Human DIT
12	327	74.0	555	5 AAU10794	Aau10794 Human zif
13	327	74.0	581	4 AEG22742	Abg22742 Novel hum
14	327	74.0	581	4 AEN99867	Adn99867 Novel hum
15	327	74.0	670	5 AAE25293	Aae25293 Novel hum
16	327	74.0	673	7 ADM06183	Adm06183 Human nuc
17	327	74.0	728	6 ABUL1645	Abul1645 Human MOD
18	327	74.0	905	8 ADM99866	Adm99866 Novel hum
19	327	74.0	905	8 ADM99577	Adm99577 Novel hum
20	325	73.8	211	7 ADL33678	Adl33678 Zinc fing
21	325	73.5	107	9 AEB77950	Aeb77950 Zinc fing
22	325	73.5	184	4 AAE06005	Aae06005 Zinc fing
23	320	72.6	184	4 AAE06002	Aae06002 Zinc fing
24	319	72.2	111	8 ADP70952	Adp70952 Novel zin
25	319	72.2	592	7 ADM04984	Adm04984 Human pro

25	319	72.2	622	4 AAM78947	Aam78947 Human pro
26	319	72.2	631	4 AAM79931	Aam79931 Human pro
27	319	72.2	632	4 ABG18386	Abg18386 Novel hum
28	318	71.9	363	8 ADG74180	Adg74180 Artificia
29	318	71.9	365	8 ADG74182	Adg74182 Artificia
30	318	71.9	365	8 ADG74177	Adg74177 Artificia
31	317	71.7	111	8 ADP70938	Adp70938 Novel zin
32	316	71.5	99	5 AAE21128	Aae21128 Zinc fing
33	316	71.5	99	7 AAE38667	Aae38667 Zinc fing
34	316	71.5	113	8 ADP70928	Adp70928 Novel zin
35	316	71.5	394	6 ABU96726	Abu96726 Human nuc
36	315	71.3	107	7 ADL33484	Adl33484 Artificia
37	315	71.3	173	4 AAB94061	Aab94061 Human pro
38	315	71.3	209	7 ADL33670	Adl33670 Zinc fing
39	315	71.3	412	8 ADR09901	Adr09901 Human pro
40	315	71.3	434	6 ABU96677	Abu96677 Human nuc
41	315	71.3	450	4 AAO05464	Aao05464 Human pol
42	315	71.3	499	7 ADM04269	Adm04269 Human pro
43	315	71.3	530	7 ADB65280	Adb65280 Human pro
44	315	71.3	553	7 ADC31307	Adc31307 Human nov
45	315	71.3	563	8 ADR10447	Adr10447 Human pro

ALIGNMENTS

RESULT 1
ADL33662
ID ADL33662 standard; protein; 378 AA.
XX
AC ADL33662;
XX
DT 20-MAY-2004 (first entry)
XX
DE Zinc finger domain-containing protein - Neurol-p65.
XX
KW modified cell; artificial transcription factor; stress resistance;
KW phenotypic trait alteration; zinc finger domain.
XX
OS Unidentified.
XX
PN WO2003048345-A1.
XX
PD 12-JUN-2003.
XX
PF 07-DEC-2002; 2002WO-KR002309.
XX
PR 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
XX (TOOL-) TOOLGEN INC.
XX Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
XX Jang Y;
XX WPI; 2003-513760/48.
XX N-PSDB; ADL33661.
XX
XX New modified cell comprising a heterologous nucleic acid encoding an
XX artificial transcription factor that confers stress resistance, useful
XX for altering a phenotypic trait of a cell or organism.
XX
XX Disclosure; SEQ ID NO 202; 169pp; English.
XX
XX The invention comprises a modified cell containing a heterologous nucleic
XX acid encoding an artificial transcription factor that confers stress
XX resistance to the modified cell. The modified cell of the invention is
XX useful for altering a phenotypic trait of a cell or organism. The present
XX amino acid sequence represents a protein which contains zinc finger
XX domains.

Db 91

DB	91	GRAFNRRSHLTRHRIH 107	: : :
RESULT 4			
ID	ADP70959	standard; protein; 111 AA.	
XX	AC	ADP70959;	
XX	DT	09-SEP-2004 (first entry)	
XX	DE	Novel zinc finger protein-related DNA-binding domain SeqID67.	
XX	DE	DNA binding domain; zinc finger domain; VEGF; VEGF-A;	
KW	KW	transcription activation domain; transcription repression domain;	
KW	KW	protein transduction domain; p65; VP16 activation domain; Kid;	
KW	KW	KOX repression domain; TAT protein; VP22 protein;	
KW	KW	Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;	
KW	KW	neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;	
ds.			
XX	OS	Unidentified.	
XX	XX	WO2004053130-A1.	
XX	PD	24-JUN-2004.	
XX	XX	09-DEC-2003; 2003WO-KR002693.	
XX	PF	09-DEC-2002; 2002US-0431892P.	
XX	PR	(TOOL-) TOOLGEN INC.	
XX	PA	Kim J, Shin H, Kwon H;	
XX	PI	WPI; 2004-468871/44.	
XX	DR		
XX	XX	New polypeptide comprising a DNA binding domain that includes a plurality	
PT	PT	of zinc finger domains, and at least two of the zinc finger domains each	
PT	PT	include respective zinc finger domain motifs, useful for modulating	
PT	PT	angiogenesis.	
XX	PS	Example 7; SEQ ID NO 67; 176pp; English.	
XX	CC	This invention relates to a novel polypeptide comprising a DNA binding	
CC	CC	domain that includes a plurality of zinc finger domains, where the DNA	
CC	CC	binding domain can bind to a site in a VEGF gene, and at least two of the	
CC	CC	zinc finger domains each include respective zinc finger domain motifs	
CC	CC	listed in the specification. The VEGF gene is the human VEGF-A gene. The	
CC	CC	polypeptide regulates the VEGF gene expression. The polypeptide further	
CC	CC	comprises a transcription activation domain, a transcription repression	
CC	CC	domain, or a protein transduction domain. The transcription activation	
CC	CC	domain comprises p65 or VP16 activation domain. The transcription	
CC	CC	repression domain comprises Kid or KOX repression domain. The protein	
CC	CC	transduction domain is a part of TAT protein, VP22 protein, or	
CC	CC	Antennapedia homeodomain. The invention may be useful for the production	
CC	CC	of compounds with a cytostatic or antiinflammatory activity acting as	
CC	CC	angiogenesis modulators. The polypeptides and nucleic acids are useful	
CC	CC	for modulating angiogenesis whilst the compounds may be useful for	
CC	CC	preventing or treating a neoplastic disorder, an inflammatory disorder,	
CC	CC	or an angiogenesis-based disorder. The present sequence is that of a DNA-	
CC	CC	binding domain which is related to the polypeptide of the invention.	
XX	SQ	Sequence 111 AA;	
	Query Match	75.6%; Score 334; DB 8; Length 111;	
	Best Local Similarity	75.3%; Pred. No. 1.6e-30;	
	Matches	58; Conservative 7; Mismatches 12; Indels 0; Gaps 0;	
QY	1	CKDCGKAFIQKSNLIRHQRTHGTGKPYKCECGKAFQSSNLTKKKIHGTGKPYKCKQC 60	
DB	31	CPDCGKSFQSSLLIRHQRTHGTGKPYKCECGKAFQSSNLTKKKIHGTGKPYKCKQC 90	
QY	61	GRAFGCPNLRHGRTH 77	
DB	91	GRAFNRRSHLTRHRIH 107	: : :
RESULT 5			
ID	ADP70948	standard; protein; 111 AA.	
XX	AC	ADP70948;	
XX	DT	09-SEP-2004 (first entry)	
XX	DE	Novel zinc finger protein-related DNA-binding domain SeqID56.	
XX	DE	DNA binding domain; zinc finger domain; VEGF; VEGF-A;	
KW	KW	transcription activation domain; transcription repression domain;	
KW	KW	protein transduction domain; p65; VP16 activation domain; Kid;	
KW	KW	KOX repression domain; TAT protein; VP22 protein;	
KW	KW	Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;	
KW	KW	neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;	
ds.			
XX	OS	Unidentified.	
XX	XX	WO2004053130-A1.	
XX	PD	24-JUN-2004.	
XX	XX	09-DEC-2003; 2003WO-KR002693.	
XX	PF	09-DEC-2002; 2002US-0431892P.	
XX	PR	(TOOL-) TOOLGEN INC.	
XX	PA	Kim J, Shin H, Kwon H;	
XX	PI	WPI; 2004-468871/44.	
XX	DR		
XX	XX	New polypeptide comprising a DNA binding domain that includes a plurality	
PT	PT	of zinc finger domains, and at least two of the zinc finger domains each	
PT	PT	include respective zinc finger domain motifs, useful for modulating	
PT	PT	angiogenesis.	
XX	PS	Example 7; SEQ ID NO 56; 176pp; English.	
XX	CC	This invention relates to a novel polypeptide comprising a DNA binding	
CC	CC	domain that includes a plurality of zinc finger domains, where the DNA	
CC	CC	binding domain can bind to a site in a VEGF gene, and at least two of the	
CC	CC	zinc finger domains each include respective zinc finger domain motifs	
CC	CC	listed in the specification. The VEGF gene is the human VEGF-A gene. The	
CC	CC	polypeptide regulates the VEGF gene expression. The polypeptide further	
CC	CC	comprises a transcription activation domain, a transcription repression	
CC	CC	domain, or a protein transduction domain. The transcription activation	
CC	CC	domain comprises p65 or VP16 activation domain. The transcription	
CC	CC	repression domain comprises Kid or KOX repression domain. The protein	
CC	CC	transduction domain is a part of TAT protein, VP22 protein, or	
CC	CC	Antennapedia homeodomain. The invention may be useful for the production	
CC	CC	of compounds with a cytostatic or antiinflammatory activity acting as	
CC	CC	angiogenesis modulators. The polypeptides and nucleic acids are useful	
CC	CC	for modulating angiogenesis whilst the compounds may be useful for	
CC	CC	preventing or treating a neoplastic disorder, an inflammatory disorder,	
CC	CC	or an angiogenesis-based disorder. The present sequence is that of a DNA-	
CC	CC	binding domain which is related to the polypeptide of the invention.	
XX	SQ	Sequence 111 AA;	
	Query Match	74.7%; Score 330; DB 8; Length 111;	
	Best Local Similarity	71.4%; Pred. No. 4.8e-30;	
	Matches	55; Conservative 11; Mismatches 11; Indels 0; Gaps 0;	
QY	1	CKDCGKAFIQKSNLIRHQRTHGTGKPYKCECGKAFQSSNLTKKKIHGTGKPYKCKQC 60	
DB	31	CRKCGRGSFQSSLLIRHQRTHGTGKPYKCECGKAFQSSNLTKKKIHGTGKPYKCKQC 90	
QY	31	CRKCGRGSFQSSLLIRHQRTHGTGKPYKCECGKAFQSSNLTKKKIHGTGKPYKCKQC 90	

QY 61 GKAFGCPNLRHGRTH 77
 DB 91 GKAFRWPENLTKRKH 107

RESULT 6
 ADP70930
 ID ADP70930 standard; protein; 111 AA.
 XX AC
 XX ADP70930;
 XX DT
 XX 09-SEP-2004 (first entry)
 XX DE
 XX Novel zinc finger protein-related DNA-binding domain SeqID38.
 XX DE
 XX DNA binding domain; zinc finger domain; VEGF; VEGF-A;
 KW transcription activation domain; transcription repression domain;
 KW protein transduction domain; p65; VP16 activation domain; Kid;
 KW KRX repression domain; TAT protein; VP22 protein;
 KW Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;
 KW neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;
 KW ds.
 KW XX
 OS Unidentified.
 XX OS
 XX WO2004053130-A1.
 XX PN
 XX 24-JUN-2004.
 XX PD
 XX 09-DEC-2003; 2003WO-KR002693.
 XX PF
 XX 09-DEC-2002; 2002US-0431892P.
 XX PR
 XX (TOOL-) TOOLGEN INC.
 XX PA
 XX Kim J, Shin H, Kwon H;
 XX PI
 XX WPI; 2004-468871/44.
 XX DR
 XX New polypeptide comprising a DNA binding domain that includes a plurality
 PT of zinc finger domains, and at least two of the zinc finger domains each
 PT include respective zinc finger domain motifs, useful for modulating
 PT angiogenesis.
 XX XX
 PS Example 7; SEQ ID NO 38; 176pp; English.
 CC This invention relates to a novel polypeptide comprising a DNA binding
 CC domain that includes a plurality of zinc finger domains, where the DNA
 CC binding domain can bind to a site in a VEGF gene, and at least two of the
 CC zinc finger domains each include respective zinc finger domain motifs
 CC listed in the specification. The VEGF gene is the human VEGF-A gene. The
 CC polypeptide regulates the VEGF gene expression. The polypeptide further
 CC comprises a transcription activation domain, a transcription repression
 CC domain, or a protein transduction domain. The transcription activation
 CC domain comprises p65 or VP16 activation domain. The transcription
 CC repression domain comprises Kid or KRX repression domain. The protein
 CC transduction domain is a part of TAT protein, VP22 protein, or
 CC Antennapedia homeodomain. The invention may be useful for the production
 CC of compounds with a cytosolic or antiinflammatory activity acting as
 CC angiogenesis modulators. The polypeptides and nucleic acids are useful
 CC for modulating angiogenesis whilst the compounds may be useful for
 CC preventing or treating a neoplastic disorder, an inflammatory disorder,
 CC or an angiogenesis-based disorder. The present sequence is that of a DNA-
 CC binding domain which is related to the polypeptide of the invention.
 XX XX
 SQ Sequence 111 AA;

Query Match 74.2%; Score 328; DB 8; Length 111;
 Best Local Similarity 71.4%; Pred. No. 8.1e-30;
 Matches 55; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 CKDCKGAFKQSNLRHGRTHGKPKCEGKAPTSNLTGKHKIHTGKPKYCKQC 60

Best Local Similarity	75.3†; Pred. No. 8.3e-30;
Matches	58; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
Qy	1 CKDCGKAFIQKSNLIRHQRTHTEKPKYCECGKAFQSSNLTGVKHKIHTGKPKYCKQC 60
	: : : :
Dd	33 CPDCGKFSQSSNLIHQHRTHTGKPKYCECGKAFQSSHLTHKLIHTGKPYICRK 92
Qy	61 GKAFGCPSNLRRHGRTH 77
	:
Dd	93 GRGFSRKSNLIRHORTH 109

[illegible]

The invention relates to a novel method for regulating a target gene. The invention further comprises: a composition for regulating a target gene, the composition comprising a zinc finger protein that can bind to the target gene or a nucleic acid encoding the protein; and a RNA molecule that includes a strand including a sequence complementary to an mRNA transcribed from the target gene; treating a neoplastic disorder in a subject comprising providing the composition, and administering the composition to the subject to modulate VEGF-A gene expression in a cell of the subject; a kit that comprises the zinc finger protein, and the RNA molecule; a modified eukaryotic cell that comprises the zinc finger protein, and the RNA molecule; and modulating expression of a target gene in a cell of a subject by administering a first component that comprises a zinc finger protein or a nucleic acid encoding the zinc finger protein, and a second component that comprises a RNA molecule or a nucleic acid that can produce the RNA molecule, to the subject, where the zinc finger protein can bind to the target gene and modulate transcription of the target gene, the RNA molecule can reduce translatability of a transcript of the target gene, and the first and second components are administered in amounts effective to modulate expression of the target gene in a cell of the subject. The method of the invention has cytostatic activity. The method is useful for regulating a target gene at both transcription and post-transcription levels. This sequence represents a human VEGF-A gene targeting zinc finger protein of the invention.

Query Match 74.2%; Score 328; DB 8; Length 113;

Sequence 125 AA:

Query Match 74.2%; Score 328; DB 8; Length 125;
Best Local Similarity 75.3%; Pred. No. 9.3e-30;
Matches 58; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 60
Db 42 CPDCKSFQSQSSLLIRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 101
QY 61 GKAFGCPNLRHGRTH 77
Db 102 GRGFSRKSNLIRHQRTH 118

RESULT 10
ABG59963
ID ABG59963 standard; protein; 213 AA.

AC ABG59963;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human DITHP polypeptide #21.

XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KW inflammatory disorder; viral infection; bacterial infection; seizure;
KW fungal infection; parasitic infections; developmental disorder; breast;
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KW thymus.

XX Homo sapiens.

OS WO200220754-A2.

PN 14-MAR-2002.

PD 29-AUG-2001; 2001WO-US027127.

XX 05-SEP-2000; 2000US-0229747P.

PR 05-SEP-2000; 2000US-0229749P.

PR 05-SEP-2000; 2000US-0229749P.

PR 05-SEP-2000; 2000US-0229750P.

PR 05-SEP-2000; 2000US-0229751P.

PR 06-SEP-2000; 2000US-0230583P.

PR 06-SEP-2000; 2000US-0230514P.

PR 06-SEP-2000; 2000US-0230515P.

PR 06-SEP-2000; 2000US-0230517P.

PR 06-SEP-2000; 2000US-0230518P.

PR 06-SEP-2000; 2000US-0230519P.

PR 06-SEP-2000; 2000US-0230595P.

PR 06-SEP-2000; 2000US-0230597P.

PR 06-SEP-2000; 2000US-0230598P.

PR 06-SEP-2000; 2000US-0230599P.

PR 06-SEP-2000; 2000US-0230610P.

PR 06-SEP-2000; 2000US-0230865P.

PR 07-SEP-2000; 2000US-0230989P.

PR 07-SEP-2000; 2000US-0230951P.

PR 07-SEP-2000; 2000US-0231163P.

PR 07-SEP-2000; 2000US-0231167P.

XX (INCY-) INCYTE GENOMICS INC.

XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;

PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;

PI Momiyama MG, Bradley DB, Kohatgi SD, Harris B, Roseberry AM;

PI Gerstin EH, Feralta CH, David MH, Panzer SR, Flores V, Daffo A;

PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;

XX WPI; 2002-383054/41.

DR N-PSDB; ABK71555.

XX An isolated polynucleotide useful in diagnostics and therapeutics.

XX Claim 29; Page 543; 686pp; English.

XX The invention relates to human diagnostic and therapeutic (dithp)
CC polynucleotides and their associated polypeptides (DITHP polypeptides).
CC The sequences of the invention are used in the treatment and diagnosis of
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
CC infections, parasitic infections, developmental disorders (e.g. anaemia,
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
CC ABG59943-ABG50220 represent human DITHP polypeptides of the invention
XX
SQ Sequence 213 AA;

Query Match 74.0%; Score 327; DB 5; Length 213;

Best Local Similarity 72.7%; Pred. No. 2.2e-29;

Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 60

Db 104 CKQCGKAFGSAASHLQMGHRTHTGKPYECKQCGSFGCASRLQMGHRTHTGKPKYCKQC 163

QY 61 GKAFGCPNLRHGRTH 77

Db 164 GKAFGCPNLRHGRTH 180

RESULT 11

AAU10794

ID AAU10794 standard; protein; 483 AA.

XX AAU10794;

DT 26-MAR-2002 (first entry)

XX Human zinc finger protein 53.

XX Human; zinc finger protein 53; malignant tumour; HIV infection;
immunological disease; inflammation.

OS Homo sapiens.

XX CN1307009-A.

PD 08-AUG-2001.

PF 21-JAN-2000; 2000CN-00111472.

PR 21-JAN-2000; 2000CN-00111472.

PA (BODA-) BODAO GENE TECH CO LTD SHANGHAI.

PI Mao Y, Xie Y;

XX WPI; 2002-049871/07.

DR N-PSDB; AAS18786.

XX Polypeptide-human zinc finger protein 53 and polynucleotide for coding
PT said polypeptide.

XX Claim 1; Page 26-27 (Disclosure); 37pp; Chinese.

PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463709P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471338P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476641P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486468P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
DR N-PSDB; ADN99083.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 14; SEQ ID NO 1467; 532pp; English.
PS
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
XX Sequence 581 AA;
SQ
Query Match 74.0%; Score 327; DB 8; Length 581;
Best Local Similarity 72.7%; Pred. No. 6.7e-29;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 1 CKDCGKAFIQSNLIRHQRHTGTGKPYKCECGKAPTQSSNLTGHKHTGTGKPYKCKQC 60

Db 472 CKQCGKAFGSGASHLQMHGRHTGTGKPYKCECKGSGFCASRLQMHGRHTGTGKPYKCKQC 531
QY 61 GKAFGCPNSLRRHQRTH 77
Db 532 GKAFGCPNSLRRHQRTH 548
RESULT 14
AAE25293
ID AAE25293 standard; protein; 670 AA.
XX
XX AAE25293;
XX
XX 30-OCT-2002 (first entry)
XX Human nucleic acid-associated protein (NAAP-12).
XX Human; nucleic acid-associated protein; NAAP-12; neurological disorder;
KW arteriosclerosis; cancer; cell proliferative disorder; atherosclerosis;
KW lymphoma; epilepsy; Alzheimer's disease; developmental; anticonvulsant;
KW autoimmune disorder; AIDS; allergy; anaemia; stroke; malaria; leishmania;
KW gene therapy; nootropic; neuroprotective; cerebroprotective; virucide;
KW immunosuppressive; protozoacide; antimicrobial.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 1. 75
FT /note= "KRAB BOX domain"
FT Region 4. 74
FT /note= "KRAB BOX"
XX WO200250279-A2.
XX
XX 27-JUN-2002.
XX
XX 19-DEC-2001; 2001WO-US050256.
XX
XX 21-DEC-2000; 2000US-0257714P.
PR 05-JAN-2001; 2001US-0260081P.
PR 16-JAN-2001; 2001US-0262302P.
PR 23-JAN-2001; 2001US-0263823P.
PR 02-FEB-2001; 2001US-0266088P.
PR 29-OCT-2001; 2001US-0348442P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughin MR, Lu Y, Arvizu C, Ramkumar J, Yao MG, Policky JL;
PI Walia NK, Tribouley KM, Yue H, Batra S, Ding L, Lal PG;
PI Borowsky ML, Lu DM, Gandhi AR, Griffin JA, Xu Y, Azimzai Y;
PI Gietzen KU, Tang YT, Warren BA, Mason PM, Burford N, Haralia AJA;
PI Lee EA, Yang J, Gorvad AE, Emerling BM, Marquis JP, Lee SY;
PI Swarnakar A, Reddy R;
XX
XX WPI; 2002-519887/55.
DR N-PSDB; AAD41202.
XX
XX Nucleic acid associated proteins and nucleic acids for diagnosing,
PT treating and preventing cell proliferative (e.g. cancers), neurological
PT (e.g. epilepsy or stroke) or autoimmune disorders (e.g. AIDS).
XX
XX Claim 67; Page 167-169; 193pp; English.
XX
XX The invention relates to nucleic acid-associated proteins (NAAP) and
CC nucleic acids. The nucleic acid and amino acid sequences are useful for
CC diagnosing, treating and preventing cell proliferative e.g.
CC arteriosclerosis, atherosclerosis, lymphoma or cancer), neurological
CC (e.g. epilepsy, Alzheimer's disease or stroke), developmental, and
CC autoimmune disorders (e.g. AIDS, allergies, or anaemia) or infections
CC (e.g. malaria, or leishmania), as well as in assessing the effects of
CC exogenous compound on the expression of nucleic acid and amino acid
CC sequences of nucleic acid-associated proteins. The invention is useful in

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:02:57 ; Search time 45 Seconds
(without alignments)
141.467 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFIQSNLRHQRTH.....KQCGKAFGCPNLRHRGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/aaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/aaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/aaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/aaa/PCUTUS COMB.pep.*
5: /cgn2_6/ptodata/1/aaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	71.3	530	2	US-10-104-047-3434
2	315	71.3	656	2	US-09-949-016-10075
3	315	71.3	1191	2	US-09-949-016-6356
4	314	71.0	292	2	US-09-538-092-1334
5	309	69.9	321	2	US-09-538-092-1304
6	309	69.9	321	2	US-09-949-016-6364
7	309	69.9	334	2	US-09-949-016-8516
8	309	69.9	409	2	US-10-104-047-2663
9	309	69.9	475	2	US-10-104-047-3470
10	308	69.7	540	2	US-10-104-047-3748
11	306	69.2	286	2	US-10-104-047-2854
12	306	69.2	803	2	US-09-538-092-1026
13	305	69.0	586	2	US-10-104-047-2592
14	302	68.3	185	2	US-09-494-190-125
15	301	68.1	504	2	US-10-104-047-2053
16	301	68.1	575	2	US-10-104-047-3423
17	300	67.9	185	2	US-09-494-190-126
18	299	67.6	325	2	US-09-538-092-1386
19	299	67.6	350	2	US-10-104-047-2500
20	299	67.6	469	2	US-09-538-092-1195
21	299	67.6	746	2	US-09-949-016-11494
22	298	67.4	685	2	US-09-949-016-8627
23	297	67.2	487	2	US-10-104-047-2868
24	293	66.3	367	2	US-09-949-016-6722
25	293	66.3	368	1	US-08-933-750C-14
26	293	66.3	368	2	US-09-234-613-14
27	293	66.3	368	2	US-09-949-016-6775

28	293	66.3	370	2	US-09-949-016-7626	Sequence 7626, Ap
29	293	66.3	556	2	US-10-104-047-3914	Sequence 3914, Ap
30	293	66.3	638	2	US-09-949-016-11391	Sequence 11391, A
31	292	66.1	357	2	US-10-104-047-2570	Sequence 2570, Ap
32	290	65.6	307	2	US-10-104-047-2986	Sequence 2986, Ap
33	290	65.6	487	2	US-09-949-016-10055	Sequence 10055, A
34	289	65.4	465	2	US-10-104-047-2590	Sequence 2590, Ap
35	289	65.4	518	2	US-09-881-578A-4	Sequence 4, Appli
36	288	65.2	405	2	US-09-949-016-9688	Sequence 9688, Ap
37	288	65.2	429	2	US-09-949-016-9689	Sequence 9689, Ap
38	288	65.2	799	2	US-10-104-047-2929	Sequence 2929, Ap
39	287	64.9	578	2	US-09-949-016-6715	Sequence 6715, Ap
40	287	64.9	583	2	US-09-949-016-9840	Sequence 9840, Ap
41	287	64.9	643	2	US-10-104-047-2241	Sequence 2241, Ap
42	287	64.9	698	2	US-09-949-016-11419	Sequence 11419, A
43	287	64.9	711	1	US-08-820-170A-10	Sequence 10, Appl
44	287	64.9	711	2	US-09-055-699-10	Sequence 10, Appl
45	287	64.9	711	2	US-09-273-565-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-104-047-3434
; Sequence 3434, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3434
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3434

Query Match	71.3%	Score 315;	DB 2;	Length 530;
Best Local Similarity	71.4%	Pred No. 98-29;	Indels	0;
Matches	55;	Conservative	5;	Mismatches 17;
Qy	1	CKDCGKAFIQSNLRHQRTHGTGKPYKCEGKAFQTSSNLTGKHKIHTGKPYCKQC	60	
Db	412	CSECGKAFIQSNLRHQRTHGTGKPYKCEGKAFQTSSNLTGKHKIHTGKPYHCNC	471	
Qy	61	GKAFGCPNLRHRGRTH 77		
Db	472	GKAFGCPNLRHRGRTH 488		
RESULT 2				
US-09-949-016-10075				
; Sequence 10075, Application US/09949016				
; Patent No. 6812339				
; GENERAL INFORMATION:				
; APPLICANT: VENTER, J. Craig et al.				
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED				
; FILE REFERENCE: CLO01307				
; CURRENT APPLICATION NUMBER: US/09/949,016				
; CURRENT FILING DATE: 2000-04-14				
; PRIOR APPLICATION NUMBER: 60/241,755				
; PRIOR FILING DATE: 2000-10-20				
; PRIOR APPLICATION NUMBER: 60/237,768				
; PRIOR FILING DATE: 2000-10-03				
; PRIOR APPLICATION NUMBER: 60/231,498				
; PRIOR FILING DATE: 2000-09-08				


```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10075
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Human
.US-09-949-016-10075

Query Match          71.3%; Score 315; DB 2; Length 656;
Best Local Similarity 72.7%; Pred. No. 1.1e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 243 CKCGKAFIWSLTLRHKRIHTGKPYKCECGKAFSRSSLTQKHKIHTGKPYKCKEC 302

QY 61 GKAFGCPNLRHGRTH 77
Db 303 GKAFKHSSALAKHKIHH 319

RESULT 3
US-09-949-016-6356
; Sequence 6356, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6356
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6356

Query Match          71.3%; Score 315; DB 2; Length 1191;
Best Local Similarity 72.7%; Pred. No. 2.2e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 772 CKCGKAFIWSLTLRHKRIHTGKPYKCECGKAFSRSSLTQKHKIHTGKPYKCKEC 831

QY 61 GKAFGCPNLRHGRTH 77
Db 832 GKAFKHSSALAKHKIHH 848

RESULT 4
US-09-538-092-1334
; Sequence 1334, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/178,965

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6356
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6356

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10075
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Human
.US-09-949-016-10075

Query Match          71.3%; Score 315; DB 2; Length 656;
Best Local Similarity 72.7%; Pred. No. 1.1e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 243 CKCGKAFIWSLTLRHKRIHTGKPYKCECGKAFSRSSLTQKHKIHTGKPYKCKEC 302

QY 61 GKAFGCPNLRHGRTH 77
Db 303 GKAFKHSSALAKHKIHH 319

RESULT 3
US-09-949-016-6356
; Sequence 6356, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6356
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6356

Query Match          71.3%; Score 315; DB 2; Length 1191;
Best Local Similarity 72.7%; Pred. No. 2.2e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 772 CKCGKAFIWSLTLRHKRIHTGKPYKCECGKAFSRSSLTQKHKIHTGKPYKCKEC 831

QY 61 GKAFGCPNLRHGRTH 77
Db 832 GKAFKHSSALAKHKIHH 848

RESULT 4
US-09-538-092-1334
; Sequence 1334, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/178,965

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10075
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Human
.US-09-949-016-10075

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10075
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Human
.US-09-949-016-10075

Query Match          71.0%; Score 314; DB 2; Length 292;
Best Local Similarity 70.1%; Pred. No. 6.1e-29;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 102 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKEC 161

QY 61 GKAFGCPNLRHGRTH 77
Db 162 GKAFGQKKYLKQKNIH 178

RESULT 5
US-09-538-092-1304
; Sequence 1304, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 1304
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q13360
US-09-538-092-1304

Query Match          69.9%; Score 309; DB 2; Length 321;
Best Local Similarity 66.2%; Pred. No. 2.6e-28;
Matches 51; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 126 CSDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKEC 185

QY 61 GKAFGCPNLRHGRTH 77
Db 186 GKFTVFSLLQKVRTH 202

RESULT 6
US-09-949-016-6364
; Sequence 6364, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6364
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-6364

Query Match      69.9%; Score 309; DB 2; Length 321;
Best Local Similarity 66.2%; Pred. No. 2.7e-28;
Matches 51; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 126 CSDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 185

QY 61 GKAFGCPNLRHGRTH 77
Db 186 GKAFVPSLQKHVRTH 202

RESULT 7
US-09-949-016-8516
; Sequence 8516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8516
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8516

Query Match      69.9%; Score 309; DB 2; Length 334;
Best Local Similarity 66.2%; Pred. No. 2.7e-28;
Matches 51; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 139 CSDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 198

QY 61 GKAFGCPNLRHGRTH 77
Db 199 GKAFVPSLQKHVRTH 215

RESULT 8
US-10-104-047-2663
; Sequence 2663, Application US/10104047
; Patent No. 6943241

```

```

; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2663
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2663

Query Match      69.9%; Score 309; DB 2; Length 409;
Best Local Similarity 67.5%; Pred. No. 3.4e-28;
Matches 52; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 247 CKECGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 306

QY 61 GKAFGCPNLRHGRTH 77
Db 307 GKAFQSSQLRQHQRIH 323

RESULT 9
US-10-104-047-3470
; Sequence 3470, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3470
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3470

Query Match      69.9%; Score 309; DB 2; Length 475;
Best Local Similarity 67.5%; Pred. No. 4e-28;
Matches 52; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 312 CKECGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 371

QY 61 GKAFGCPNLRHGRTH 77
Db 372 GKAFQSSQLRQHQRIH 388

RESULT 10
US-10-104-047-3748
; Sequence 3748, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:

```

```
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3748
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3748

Query Match      69.7%; Score 308; DB 2; Length 540;
Best Local Similarity 67.5%; Pred. No. 6.1e-28;
Matches 52; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 286 CEECGKAFQTSSTLTGKRIHTGKPYKCEGCGKAFNRSSKLTGKHNHTGKPYKCEC 345

Qy 61 GKAFGCPNLRHGRTH 77
Db 346 GKAFNRSSNLTGKRIH 362

RESULT 11
US-10-104-047-2854
; Sequence 2854, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2854
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2854

Query Match      69.2%; Score 306; DB 2; Length 286;
Best Local Similarity 64.9%; Pred. No. 5.2e-28;
Matches 50; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 40 CKQCGKAFSWHSSVRIHERHTGKPYKCEGCGKSNFSSFRHRHTGKPYKCEC 99

Qy 61 GKAFGCPNLRHGRTH 77
Db 100 GKAFNCPSSFHRHERTH 116

RESULT 12
US-09-538-092-1026
; Sequence 1026, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapPatSeqFormatter Version 0.9
; SEQ ID NO 1026

; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3748
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3748

Query Match      69.7%; Score 308; DB 2; Length 540;
Best Local Similarity 67.5%; Pred. No. 6.1e-28;
Matches 52; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 286 CEECGKAFQTSSTLTGKRIHTGKPYKCEGCGKAFNRSSKLTGKHNHTGKPYKCEC 345

Qy 61 GKAFGCPNLRHGRTH 77
Db 346 GKAFNRSSNLTGKRIH 362

RESULT 11
US-10-104-047-2854
; Sequence 2854, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2854
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2854

Query Match      69.2%; Score 306; DB 2; Length 286;
Best Local Similarity 64.9%; Pred. No. 5.2e-28;
Matches 50; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 40 CKQCGKAFSWHSSVRIHERHTGKPYKCEGCGKSNFSSFRHRHTGKPYKCEC 99

Qy 61 GKAFGCPNLRHGRTH 77
Db 100 GKAFNCPSSFHRHERTH 116

RESULT 12
US-09-538-092-1026
; Sequence 1026, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapPatSeqFormatter Version 0.9
; SEQ ID NO 1026

; LENGTH: 803
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P28160
US-09-538-092-1026

Query Match      69.2%; Score 306; DB 2; Length 803;
Best Local Similarity 71.4%; Pred. No. 1.6e-27;
Matches 55; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 534 CEECGKAFNFSILTGKRIHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCEC 593

Qy 61 GKAFGCPNLRHGRTH 77
Db 594 GKAFQTQSSNLTGKRIH 610

RESULT 13
US-10-104-047-2592
; Sequence 2592, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2592
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2592

Query Match      69.0%; Score 305; DB 2; Length 586;
Best Local Similarity 68.8%; Pred. No. 1.5e-27;
Matches 53; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 371 CDECGKAFNQSSLTGKRIHTGKPYKCEGCGKAFQSSLTGKHKIHTGKPYKCEC 430

Qy 61 GKAFGCPNLRHGRTH 77
Db 431 GKAFSWSSAFTGKRIH 447

RESULT 14
US-09-494-190-125
; Sequence 125, Application US/09494190
; Patent No. 6610512
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: TSRI 645.2
; CURRENT APPLICATION NUMBER: US/09/494,190
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: EP/99/07742
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/173,941
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 125
; LENGTH: 185
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:expressed
US-09-494-190-125

Query Match      68.3%; Score 302; DB 2; Length 185;
Best Local Similarity 63.6%; Pred. No. 9.4e-28;
Matches 49; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFTQKSNLIRHQRTHTGKPYKCECGKAFQTQSSNLTKKKKIHTGKPYKCKQC 60
   |::|||::|:|:|||||::|:|||||::|:|||||::|:|||||::|:|||||::|:|
Db 15 CPEGKSFGRKDSLVRHQRTHTGKPYKCECGKAFQTQSSNLTKKKKIHTGKPYKCKQC 74
   |::|||::|:|:|||||::|:|||||::|:|||||::|:|||||::|:|||||::|:|

QY 61 GKAFGCPNLRHGRTH 77
   |||:|:|:|:|:|
Db 75 GKSFSDCRDLARHQRTH 91
   |||:|:|:|:|
```

```
RESULT 15
US-10-104-047-2053
; Sequence 2053, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2053
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2053
```

```
Query Match      68.1%; Score 301; DB 2; Length 504;
Best Local Similarity 67.5%; Pred. No. 3.7e-27;
Matches 52; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 CKDCGKAFTQKSNLIRHQRTHTGKPYKCECGKAFQTQSSNLTKKKKIHTGKPYKCKQC 60
   |::|||::|:|:|||||::|:|||||::|:|||||::|:|||||::|:|||||::|:|
Db 395 CEECGSAFRAFSTLTGKRVHTGKPYKCECGKAFNWSSTLTKKRIHTGKPYKCEC 454
   |::|||::|:|:|||||::|:|||||::|:|||||::|:|||||::|:|||||::|:|

QY 61 GKAFGCPNLRHGRTH 77
   |||:|:|:|:|
Db 455 GKAFNRSSNLTTRHKTIH 471
   |||:|:|:|
```

Search completed: January 9, 2006, 20:10:34
Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:03:37 ; Search time 61 Seconds
(without alignments)
527.424 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCGKAFIQKSLNRHQT.....KQCGKAFGCPNLRHGRTH 77

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	378	4	US-10-314-669-2
2	442	100.0	378	4	US-10-669-861-2
3	374	84.6	209	4	US-10-314-669-14
4	374	84.6	209	4	US-10-669-861-14
5	340	76.9	107	5	US-10-746-864-60
6	340	76.9	111	5	US-10-732-620-57
7	334	75.6	111	5	US-10-732-620-67
8	330	74.7	111	5	US-10-732-620-56
9	328	74.2	111	5	US-10-732-620-38
10	328	74.2	111	5	US-10-732-620-58
11	328	74.2	113	5	US-10-732-620-44
12	327	74.0	555	5	US-10-450-763-53101
13	327	74.0	673	4	US-10-108-260A-4868
14	326	73.8	211	4	US-10-314-669-264
15	325	73.5	107	5	US-10-746-864-54
16	325	73.5	184	3	US-09-765-555-42
17	321	72.6	184	3	US-09-765-555-39
18	320	72.4	111	5	US-10-732-620-60
19	319	72.2	592	4	US-10-108-260A-3669
20	319	72.2	632	5	US-10-450-763-48745
21	317	71.7	111	5	US-10-732-620-46
22	316	71.5	99	4	US-09-911-261A-8
23	316	71.5	99	4	US-10-057-408-8
24	316	71.5	99	4	US-10-333-487-8
25	316	71.5	113	5	US-10-732-620-36
26	315	71.3	107	4	US-10-314-669-23
27	315	71.3	107	4	US-10-669-861-23

28	315	71.3	209	4	US-10-314-669-10	Sequence 10, Appl
29	315	71.3	209	4	US-10-669-861-10	Sequence 10, Appl
30	315	71.3	499	4	US-10-108-260A-2954	Sequence 2954, Ap
31	315	71.3	530	4	US-10-104-047-3434	Sequence 3434, Ap
32	315	71.3	1191	4	US-10-408-765A-2105	Sequence 2105, Ap
33	315	71.3	1212	5	US-10-450-763-30758	Sequence 30758, A
34	314	71.0	111	5	US-10-732-620-33	Sequence 33, Appl
35	314	71.0	111	5	US-10-732-620-39	Sequence 39, Appl
36	314	71.0	111	5	US-10-732-620-62	Sequence 62, Appl
37	312	70.6	183	5	US-10-732-620-27	Sequence 27, Appl
38	312	70.6	111	5	US-10-732-620-59	Sequence 59, Appl
39	312	70.6	184	3	US-09-765-555-38	Sequence 38, Appl
40	312	70.6	184	3	US-09-765-555-75	Sequence 75, Appl
41	311	70.4	99	3	US-09-911-261A-9	Sequence 9, Appl
42	311	70.4	99	3	US-09-911-261A-10	Sequence 10, Appl
43	311	70.4	99	4	US-10-057-408-9	Sequence 9, Appl
44	311	70.4	99	4	US-10-057-408-10	Sequence 10, Appl
45	311	70.4	99	4	US-10-333-487-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-10-314-669-2
; Sequence 2, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-2

Query Match	100.0%;	Score 442;	DB 4;	Length 378;
Best Local Similarity	100.0%;	Pred. No. 9.8e-39;		
Matches	77;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	CKDCGKAFIQKSLNRHQTHTGEPYKCEGCKAFTOSSNLTKHKIKHTGSKPYKCKQC	60	
Db	33	CKDCGKAFIQKSLNRHQTHTGEPYKCEGCKAFTOSSNLTKHKIKHTGSKPYKCKQC	92	
Qy	61	GKAFGCPNLRHGRTH	77	
Db	93	GKAFGCPNLRHGRTH	109	
RESULT 2				

7


```

; TITLE OF INVENTION: REGULATION OF PROKARYOTIC GENE
; FILE REFERENCE: 12279-012001
; CURRENT APPLICATION NUMBER: US/10/746,864
; CURRENT FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 10/669,861
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 164
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-746-864-60

```

```

Query Match 76.9%; Score 340; DB 5; Length 107;
Best Local Similarity 74.0%; Pred. No. 1.9e-28;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 31 CDHCGKAFSVSSNLRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFGCPNLRHRGRTH 107

```

```

RESULT 6
US-10-732-620-57
; Sequence 57, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-57

```

```

Query Match 76.9%; Score 340; DB 5; Length 111;
Best Local Similarity 74.0%; Pred. No. 1.9e-28;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 3 CKTCQKFSRSDHLKTHTRHTGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 62
QY 61 GKAFGCPNLRHRGRTH 77
Db 63 GKAFGCPNLRHRGRTH 79

```

```

RESULT 7
US-10-732-620-67
; Sequence 67, Application US/10732620
; Publication No. US20050032186A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-67

```

```

Query Match 75.6%; Score 334; DB 5; Length 111;
Best Local Similarity 75.3%; Pred. No. 8.4e-28;
Matches 58; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 31 CPDCGKFSQSSSLRHRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFNRRSHLTRHQRTH 107

```

```

RESULT 8
US-10-732-620-56
; Sequence 56, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-56

```

```

Query Match 74.7%; Score 330; DB 5; Length 111;
Best Local Similarity 71.4%; Pred. No. 2.2e-27;
Matches 55; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 60
Db 31 CRKCGRFSRKSRLRHRHQRTHGTGKPYKCEGCKAFTQSSNLTGKHKIHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFNRRSHLTRHQRTH 107

```

```

RESULT 9
US-10-732-620-38
; Sequence 38, Application US/10732620

```

; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-38

Query Match 74.2%; Score 328; DB 5; Length 111;
Best Local Similarity 71.4%; Pred. No. 3.6e-27;
Matches 55; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTGKHKIHTGKPYKCKQC 60
Db 31 CSDCGKAFRDSCLNRRHRTHTGKPYKCECGKAFQSSNLTGKHKIHTGKPYKCKQC 90
Qy 61 GKAFGCPNSLRRHGRTH 77
Db 91 GKAFGCPNSLRRHGRTH 107

RESULT 10
US-10-732-620-58
; Sequence 58, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-58

Query Match 74.2%; Score 328; DB 5; Length 111;
Best Local Similarity 55.2%; Pred. No. 3.6e-27;
Matches 58; Conservative 10; Mismatches 9; Indels 28; Gaps 1;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKA----- 35
Db 3 CMECGKAFNRSHLRHQRTHTGKPYKCKQCGKAFGCPNSLRRHGRTHTGKPFQKTC 62
Qy 36 ---FTQSSNLTGKHKIHTGKPYKCKQCGKAFGCPNSLRRHGRTH 77
Db 63 QRKFSRSHLKTHTHTGKPYKCKQCGKAFGCPNSLRRHGRTH 107

RESULT 11
US-10-732-620-44

; Sequence 44, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-44

Query Match 74.2%; Score 328; DB 5; Length 113;
Best Local Similarity 75.3%; Pred. No. 3.7e-27;
Matches 58; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTGKHKIHTGKPYKCKQC 60
Db 33 CPDCGKFSOSSLIRHQRTHTGKPYKCECGKAFQSSHLATTHKIHTGKPYCKKC 92
Qy 61 GKAFGCPNSLRRHGRTH 77
Db 93 GRGFSRKSNIIRHQRTH 109

RESULT 12
US-10-450-763-53101
; Sequence 53101, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53101
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (457)..(510)
; OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00738H, p-value=2.301e-
; OTHER INFORMATION: 17, raw score of 23.08
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (85)..(437)
; OTHER INFORMATION: zinc finger, C2H2 type domain identified by Pfam, accession
; OTHER INFORMATION: name zf-C2H2, E-value=3.9e-42, Pfam score of 153.4
US-10-450-763-53101

Query Match 74.0%; Score 327; DB 5; Length 555;
Best Local Similarity 72.7%; Pred. No. 2.4e-26;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

This Page Blank (uspto)


```

; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: sigma_Eflczf2-czf3
; US-11-196-670-64

```

```

Query Match      31.7%; Score 140; DB 7; Length 54;
Best Local Similarity 51.1%; Pred. No. 7e-10;
Matches 23; Conservative 9; Mismatches 13; Indels

QY 1 CKDCGKAFIQKSNLIRHQHTGTGPKYKCEGCKAFTQSSNLTKH 45
db CGICCKKAFPKHKLHLEHMLKHSGERPKYCDKCKGKFSHGVSQH 45

```

```

RESULT 13
US-11-063-343-21
; Sequence 21, Application US/11063343
; Publication No. US20050272061A1
; GENERAL INFORMATION:
; APPLICANT: Petroziello, Joseph M.
; APPLICANT: Carter, Paul
; TITLE OF INVENTION: Expression Profiling
; TITLE OF INVENTION: Lung Cancer
; FILE REFERENCE: 2681-1-003N
; CURRENT APPLICATION NUMBER: US/11/063,343
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: 60/546,019
; PRIOR FILING DATE: 2004-02-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-063-343-21

```

```

Query Match      30.1%; Score 133; DB 7; Length 644;
Best Local Similarity 29.5%; Pred. No. 5.7e-08;
Matches 28; Conservative 14; Mismatches 33; Indels 20; Gaps 2;

Qy 1 CKDCGKAFTQKSNLRH-----QRTHTGKPYKCECGKAPTQSNL 42
Db 536 CPNCGQRFETENLVHEHSSCLDQDMFKSAIMEENERDH--RRKHFCNLCKGKGYQRC 593

Qy 43 TKHKKLIHTGKPYKCKCGKAFGCPFSLRRHGRTH 77
Db 594 REHYTVHTTKKQFVCOTCGHQFLREROLRLHNDMH 628

```

RESULT 14
US-11-196-670-61
; Sequence 61, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger Transcription Factors
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582

```

; PRIOR FILING DATE: 2000-06-09
;
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
;
; OTHER INFORMATION: SIPinzf3+nzf4
US-11-196-670-61

```

```

Query Match      29.6%; Score 131; DB 7; Length 52;
Best Local Similarity 48.9%; Pred. No. 7e-09;
Matches 22; Conservative 7; Mismatches 16; Indels

QY 1 CKDCGKAFIQKSLNLRHORTHGKPYKCECGKAFQSSNLTKH 45
      | | | | | | | | | | | | | | | | | | | |
Db 1 CTECGKAPKYHLLKHLRIHSGKPYPCNCKKRFHSGSYSH 45
      | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 15
US-11-196-670-63
; Sequence 63, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: sigma-EFLnzf3+nzf4
US-11-196-670-63

```

Query Match 29.6%; Score 131; DB 7; Length 52;
Best Local Similarity 48.9%; Pred.No. 7e-09;
Matches 22; Conservative 7; Mismatches 16; Indels

QY 1 CKDCGKAFIQKSNLRHQRTHTEGPKEPCCECGKAFTOSNLTKH 45
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1 CTECGKAFQKHLLKEHLRIHSRGPEPCPNCKRFSHGVSYSH 45

Search completed: January 9, 2006, 20:12:06
Job time : 8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:00:51 ; Search time 39 Seconds
(without alignments)
189.966 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFIQKSLIRHQT.....KQCGKAFGCPNLRHGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	71.3	1191	2 S35305	zinc finger protei
2	314	71.0	292	2 S43826	finger protein OZF
3	311	70.4	169	2 A39240	finger protein mfg
4	310	70.1	427	2 A35659	krueppel-related p
5	307	69.5	595	2 G02075	transcription repr
6	306	69.2	229	2 A48927	Kruppel-like zinc
7	306	69.2	732	2 S47073	finger protein HZF
8	306	69.2	803	2 S26823	zinc finger protei
9	301	68.1	105	2 B48827	zinc finger protei
10	301	68.1	543	2 B34612	zinc finger protei
11	299	67.6	247	2 S47070	zinc finger protei
12	299	67.6	325	2 I38616	zinc finger protei
13	299	67.6	469	2 I38600	zinc finger protei
14	299	67.6	651	2 B32891	finger protein 2
15	299	67.6	686	2 A34612	zinc finger protei
16	295	66.7	519	2 A38073	transcription acti
17	294	66.5	488	2 S47072	finger protein HZF
18	294	66.5	555	2 A56560	zinc finger protei
19	294	66.5	564	2 T12489	hypothetical prote
20	293	66.3	319	2 T46469	hypothetical prote
21	293	66.3	347	2 S00549	developmental cont
22	292	66.1	93	2 PQ0636	zinc finger protei
23	292	66.1	614	2 JH0500	zinc finger protei
24	290	65.6	109	2 H45193	zinc finger protei
25	290	65.6	474	2 I54338	zinc finger protei
26	289	65.4	200	2 S47067	finger protein HZF
27	289	65.4	542	2 A54661	zinc finger protei
28	288	65.2	348	2 I38599	zinc finger protei
29	287	64.9	386	2 T12527	hypothetical prote

30	287	64.9	594	2 T12488	hypothetical prote
31	287	64.9	728	2 A48830	probable transcrip
32	286	64.7	194	2 I53859	zinc finger protei
33	285	64.5	462	2 S10397	finger protein kox
34	285	64.5	540	2 T14748	hypothetical prote
35	283	64.0	337	2 S60520	finger protein ZNF
36	283	64.0	411	2 S10245	finger protein, te
37	283	64.0	636	2 I48689	gene NK10 protein
38	283	64.0	693	2 I37570	zinc finger protei
39	282	63.8	378	2 S33994	finger protein ZNF
40	282	63.8	553	2 S22954	finger protein zfp
41	282	63.8	555	2 I53869	zinc finger protei
42	281	63.6	195	2 S00754	zinc finger protei
43	281	63.6	242	2 S02338	finger protein ZNF
44	281	63.6	261	2 S70006	finger protein zfo
45	281	63.6	393	2 JN0533	finger protein pml

ALIGNMENTS

RESULT 1

S35305
zinc finger protein ZNF91 - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: S35305
R:Beliefroid, E.J.; Marine, J.C.; Ried, T.; Lecocq, P.J.; Riviere, M.; Anemiy, C.; Pon
EMBO J. 12, 1363-1374, 1993
A>Title: Clustered organization of homologous KRAB zinc-finger genes with enhanced exp
A:Reference number: S35305; MUID:94223677; PMID:8467795
A:Accession: S35305
A:Molecule type: mRNA
A:Residues: 1-1191 <BEL>
A:Cross-references: UNIPROT:Q05481; UNIPARC:UPI000013C42A; EMBL:L11672; NID:g186773; PI
A>Note: The authors translated the codon GCA for residue 750 as Thr and GCT for residue
C:Genetics:
A:Gene: GDB:ZNF91; HPF7; HTF10
A:Cross-references: GDB:132284
A:Map position: 19p12-19p12
C:Keywords: DNA binding; zinc finger

Query Match 71.3%; Score 315; DB 2; Length 1191;
Best Local Similarity 72.7%; Pred. No. 2e-23;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSLIRHQT.....KQCGKAFGCPNLRHGRTH 77
Db 772 CKECGKAFIQKSLIRHQT.....KQCGKAFGCPNLRHGRTH 77

QY 61 GKAFGCPNLRHGRTH 77
Db 832 GKAFGCPNLRHGRTH 848

RESULT 2

S43826
finger protein OZF, Krueppel-related - human
C:Species: Homo sapiens (man)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43826
R:le Chalony, C.; Prosperi, M.T.; Haluza, R.; Aplou, F.; Dutrillaux, B.; Goubin, G.
J. Mol. Biol. 236, 395-404, 1994
A>Title: The OZF gene encodes a protein consisting essentially of zinc finger motifs.
A:Reference number: S43826; MUID:94149744; PMID:8107129
A:Accession: S43826
A:Molecule type: mRNA
A:Residues: 1-292 <LEG>
A:Cross-references: UNIPROT:Q15072; UNIPARC:UPI0000130F91; EMBL:X70394; NID:g468707; PID
C:Genetics:
A:Gene: GDB:ZNF146; OZF
A:Cross-references: GDB:230289; OMIM:601505
A:Map position: 19q13.1-19q13.1

WSTR, RSTK, HSAK

-1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
Q S N R Q Y : SEQJND02

C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger

```

Query Match      71.0%; Score 314; DB 2; Length 292;
Beat Local Similarity 70.1%; Pred. No. 7.8e-24;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

1 CKDCGKATIDENLIRHRTITGKPYKCEGCGKAFTOSSNLTKKKLHTGEKPKKQC 60
102 CKDCGKATIQSNLIRHRTITGKPYKCEGCGKAFTOSSNLTKKKLHTGEKPKKQC 161

61 GKAFCCPSNLRHGRTH 77
162 GTAFGCKYLKQHNH 178

```

RESULT 3
A39240 finger protein mfg1 - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 09-Jul-2004
C/Accession: A39240
R/Passananti, C.; Felsani, A.; Caruso, M.; Amati, P.
Proc. Natl. Acad. Sci. U.S.A. 86, 9417-9421, 1989
A/Title: Mouse genes coding for "zinc-finger"-containing proteins: characterization and
A/Reference number: A39240; MUID:90083278; PMID:2512579
A/Accession: A39240
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-169 <PAS>
A/Cross-references: UNIPROT:P16372; UNIPARC:UPI00000029AE; GB:M28513; NID:gl99136; PIDN:
C/Keywords: DNA binding; zinc finger

```

Query Match      70.4%; Score 311; DB 2; Length 169;
Best Local Similarity 68.8%; Pred. No. 9.7e-24;
Matches 53; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CKDGGKAFIPKSNLRIQRTHTGKPKYCECGGKAPTOSNLTKHKIHTGKPKYKQC 60
   ::::: :
Db 38 CEEGCKAFIRSKLSLQRTVHTGKPKYCECGGKAFNSHNSLSEHKRIHTGKPKYCEC 97
   ::::: :

Qy 61 GKAFGCFPSNLRHGRTH 77
   ::::: |
Db 98 GKAFSTSNLSEHKRIH 114
   ::::: |

```

RESULT 4
A35659
krueppel-related protein H-plk - human
C:Species: Homo sapiens (man)
C:Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 09-Jul-2004
C:Accession: A35659
R:Kato, N.; Shimotohno, K.; VanLeeuwen, D.; Cohen, M.
Mol. Cell. Biol. 10, 4401-4405, 1990
A:Title: Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger protein
A:Reference number: A35659; MUID:90318410; PMID:2115127
A:Accession: A35659
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-427 <KAT>
A:Cross-references: UNIPROT:Q02313; UNIPARC:UPI000006B47C; GB:M55422; NID:g184342; PIDN:
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match	70.1%	Score 310;	DB 2;	Length 427;
Best Local Similarity	71.4%	Pred. No. 2.7e-23;		
Matches	55;	Conservative 6;	Mismatches 16;	Indels 0; Gaps 0;
Qy	1	CKDCGKAFIQKSNLIRHQRITHTGKPYKCEGCKAFTOSSNLTKHKHITGKEKPYKQKC	60	
Db	279	CEECGKAFNQSGSTUTTKRIHSGHKPYKCEGCKAFQFSLNLTDHKKHITGKEKPYKCEC	338	
	61	GKAFGCPSNLRRRHGRTH	77	

Db 339 GKAFNQLSNLTRHKVH 355

```

RESULT 5
G02075
transcription repressor zinc finger protein 85 - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C/Accession: G02075
R/Poncelet, D.A.
submitted to the EMBL Data Library, September 1995
A/Reference number: G09169
A/Accession: G02075
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-535 <PON>
A/Cross-references: UNIPROT:Q03923; UNIPARC:UPI0000013C428; EMBL:U35376; NID
C/Genetics:
A/Gene: GDB:ZNF85
A/Cross-references: GDB:132279
A/Map position: 19p12-19p12

```

	Query Match	69.5%	Score 307;	DB 2;	Length 595;
	Best Local Similarity	66.0%	Fred.No. 7e-23;		
	Matches 53;	Conservative	Mismatches 16;	Indels 0;	Gaps 0;
Qy	1	CKDCGKAFIQKSNLIRHQRTHTGEKPYKCECGKAFQTQSSNLTKHKKIHTGEKPYKCKQC	60		
Db	232	CBECKGAFQSSNLIRKHKIHTGEKPYKCECGKAFNRFTLTTHKLIHTGEKPYKCKEC	291		
Qy	61	GKAFGCPNLRRLRHGRTH	77		
Db	292	GKAFNRSTLTTHRKIH	308		

RESULT 6

A48927

Kruppel-like zinc finger protein HKR-T1 - human

N:Alternate names: zinc finger protein kox15; zinc finger protein ZNF22

C:Species: Homo sapiens (man)

C:Date: 19-Dec-1993 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004

C:Accession: A48927; I37954; C56403; S10403

R:Wu, B.Y.; Hanley, E.W.; Turka, L.A.; Nabel, G.J.

Blood 80, 2571-2576, 1992

A:Title: Isolation of a cDNA clone encoding a zinc finger protein highly ex

A:Reference number: A48927; MUID:93043304; PMID:1421376

A:Accession: A48927

A:Molecule type: mRNA

A:Residues: 1-229 <WU1>

A:Cross-references: UNIPROT:P17026; UNIPARC:UPI00000493DA; GB:S50223; NID:G

A:Experimental source: MOLT 4 T-cells

A:Note: the authors translated the codon ATC for residue 76 as Asn, and the

A:Note: sequence modified after extraction from NCBI backbone (NCBIN:118663

R:Thiesen, H.J.

New Biol. 2, 363-374, 1990

A:Title: Multiple genes encoding zinc finger domains are expressed in human

A:Reference number: I37949; MUID:91145339; PMID:2288909

A:Accession: I37954

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 139-194 <TH1>

A:Cross-references: UNIPARC:UPI000016ABBD; EMBL:X52346; NID:g34147; PIDN:CA

A:Note: submitted to the EMBL Data Library, March 1990

R:Bray, P.; Lichter, P.; Thiesen, H.J.; Ward, D.C.; Dawid, I.B.

Proc. Natl. Acad. Sci. U.S.A. 88, 9563-9567, 1991

A:Title: Characterization and mapping of human genes encoding zinc finger p

A:Reference number: A56409; MUID:92052132; PMID:1946370

A:Accession: C56409

A:Molecule type: DNA

A:Residues: 132-209 'Y', 211-212, 'LVQE' <BRA>

A:Cross-references: UNIPARC:UPI0000178A38; GB:M771172; NID:g340429

A:Note: the nucleotide sequence and translation in GenBank entry HUMZFPAC,

A;Reference number: S10397
A;Accession: S10421
A;Molecule type: mRNA
A;Residues: 413-468 <TH1>
A;Cross-references: UNIPARC:UPI000016ABCS; EMBL:X52335; NID:g34165; PID:CAA36561.1; PID:R;Thiesen, H.J.
New Biol. 2, 363-374, 1990
A;Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A;Reference number: I37949; MUID:91145339; PMID:2288909
A;Accession: I37972
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:00:06 ; Search time 160 Seconds
(without alignments)
339.535 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFQKSNLRHQRT.....KQCGKAFGCPNLRHRGRT 77

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	74.0	638	2	Q86VX3 HUMAN
2	327	74.0	673	1	ZNA333 HUMAN
3	319	72.2	592	2	Q8N8Q4 HUMAN
4	319	72.2	616	1	ZNF93 HUMAN
5	315	71.3	173	2	Q9H9X5 HUMAN
6	315	71.3	403	2	Q59FB0 HUMAN
7	315	71.3	503	2	Q4G180 HUMAN
8	315	71.3	530	2	Q8NA64 HUMAN
9	315	71.3	563	2	Q6ZMR0 HUMAN
10	315	71.3	1191	1	ZNF91 HUMAN
11	314	71.0	292	1	OZF HUMAN
12	314	71.0	292	2	Q5RAA6 PONPY
13	314	71.0	292	2	Q5RFP4 PONPY
14	314	71.0	544	1	ZNS02 HUMAN
15	314	71.0	589	2	Q6AYV1 RAT
16	314	71.0	611	2	Q8CII6 MOUSE
17	313	70.8	292	1	OZF BOVIN
18	313	70.8	292	2	Q920Q5 MOUSE
19	313	70.8	302	2	Q8BQK6 MOUSE
20	313	70.8	584	2	Q86XU0 HUMAN
21	313	70.8	626	2	Q99KB9 MOUSE
22	313	70.8	653	2	Q5DTH6 MOUSE
23	312	70.6	395	2	Q43693 HUMAN
24	311	70.4	169	1	ZFP58 MOUSE
25	311	70.4	338	2	Q8NC79 HUMAN
26	311	70.4	530	2	Q8NEM1 HUMAN
27	311	70.4	536	2	Q96WL5 HUMAN
28	311	70.4	588	1	ZNA429 HUMAN
29	311	70.4	688	1	ZNA20 HUMAN
30	311	70.4	783	1	ZNS588 HUMAN
31	310	70.1	427	2	Q02313 HUMAN

32	310	70.1	481	2	Q4R3H8 MACFA	Q4R3H8 macaca fasc
33	310	70.1	529	2	Q5FWH5 MOUSE	Q5FWH5 mus musculus
34	310	70.1	546	2	Q7M6X4 MOUSE	Q7M6X4 mus musculus
35	309	69.9	238	2	Q8BYD9 MOUSE	Q8BYD9 mus musculus
36	309	69.9	321	1	ZNI177 HUMAN	Q13360 homo sapien
37	309	69.9	402	2	Q8NCA3 HUMAN	Q8NCA3 homo sapien
38	309	69.9	423	2	Q6POL2 HUMAN	Q6POL2 homo sapien
39	309	69.9	471	2	Q8NCF9 HUMAN	Q8NCF9 homo sapien
40	309	69.9	475	2	Q8NA42 HUMAN	Q8NA42 homo sapien
41	309	69.9	475	2	Q6X2C7 HUMAN	Q6X2C7 homo sapien
42	309	69.9	475	2	Q4R6C3 MACFA	Q4R6C2 macaca fasc
43	309	69.9	519	2	Q96M04 HUMAN	Q96M04 homo sapien
44	308	69.7	502	2	Q59F60 HUMAN	Q59F60 homo sapien
45	308	69.7	540	2	Q8N211 HUMAN	Q8N211 homo sapien

ALIGNMENTS

RESULT 1
Q86VX3 HUMAN PRELIMINARY; PRT; 638 AA.
ID Q86VX3; HUMAN
AC Q86VX3;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE ZNF433 protein.
DE ZNF433 protein.
GN Name=ZNF433;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hestek F.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schestz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schain J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC047412; AAH47412.1; -; mRNA.
DR HSSP; P08047; 1SP2.
DR Ensembl; ENSG00000197647; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 19.
DR PRINTS; PR00048; ZINCfinger.

OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=14702039; DOI=10.1038/ngl1285;
 RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hota T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki F., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi R.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 cDNAs";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK096342; BAC04764.1; -; mRNA.
 DR HSSP; P08047; 1SP2.
 DR Ensembl; ENSG0000081665; Homo sapiens.
 DR HGNC; HGNC:23779; ZNF505.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0008350; P:transcription; IEA.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR006025; Pept M. Zn_BS.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00996; zf-C2H2; 15.
 DR PRINTS; PR00048; ZINC_FINGER.
 DR ProDom; PD000003; Znf_C2H2; 15.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 16.
 DR PROSITE; PS00805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 15.
 DR PROSITE; PS01517; ZINC_FINGER_C2H2_2; 16.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 2.
 SQ SEQUENCE 592 AA; 67920 MW; D116FB9365EB4BDE CRC64;
 Query Match 72.2%; Score 319; DB 2; Length 592;
 Best Local Similarity 70.1%; Pred. No. 1.4e-25;
 Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
 OY 1 CKDCGKAFQKQNLIRHQRTHGKPKYCEGCGKAFQSSNLTGKHKIHTGKPKYCKQC 60
 Db 427 CEECGKAFNQSSLTGKHKIHTGKPKYCEGCGKAFNQSSLTGKHKIHTGKPKYCEC 486
 OY 61 GKAFGCPNLRHGRTH 77
 ||||| | : : : : |

Db 487 GKAFNQSSLTGKHKIHT 503
 RESULT 4
 ZNF93 HUMAN
 ID ZNF93 HUMAN STANDARD; PRT; 616 AA.
 AC P35789; Q9Y2N8;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Zinc finger protein 93 (Zinc finger protein HTP34) (Fragment).
 GN Name=ZNF93;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15057824; DOI=10.1038/nature02399;
 RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
 RA Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
 RA Aerle A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
 RA Caenepeel S., Carrano A.V., Caciile C., Chan Y.M., Christensen M.,
 RA Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Dettler J.C.,
 RA Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M.,
 RA Glavina T., Gomez M., Gonzales E., Groza M., Hamon N., Hawkins T.,
 RA Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,
 RA Kobayashi A., Lartionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
 RA Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J.,
 RA Nelson K., Nolan M., Ovcharenko I., Pittluck S., Pollard M.,
 RA Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J.,
 RA Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,
 RA Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
 RA Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,
 RA Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,
 RA Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
 RA Rubin E.M., Lucas S.M.;
 RT "The DNA sequence and biology of human chromosome 19.";
 RL Nature 428:529-535 (2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 1-191.
 RX MEDLINE=91219421; PubMed=2023909;
 RA Bellefroid E.J., Poncelet D.A., Lecocq P.J., Revelant O.,
 RA Martial J.A.;
 RT "The evolutionarily conserved Kruppel-associated box domain defines a
 subfamily of eukaryotic multifingered proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:3608-3612(1991).
 CC -!- FUNCTION: May be involved in transcriptional regulation.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: Expressed early during embryonic development.
 CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
 family.
 CC -!- SIMILARITY: Contains 17 C2H2-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 KRAB domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL; AC007204; AAD22981.1; -; Genomic DNA.
 DR EMBL; M61873; AAA83548.1; -; Genomic DNA.
 DR HSSP; P08047; 1SP2.
 DR Ensembl; ENSG00000081665; Homo sapiens.
 DR HGNC; HGNC:13169; ZNF93.
 DR MIM; 603975; -.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0008270; F:zinc ion binding; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001909; KRAB.

```
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 16.
DR PRINTS; PRO00048; ZINC_FINGER.
DR ProDom; PD000003; Znf_C2H2; 16.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 17.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 16.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 17.
KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
Transcription regulation; Zinc; Zinc-finger.
FT DOMAIN <1 71
FT ZN_FING 141 163
FT ZN_FING 169 191
FT ZN_FING 197 219
FT ZN_FING 225 247
FT ZN_FING 253 275
FT ZN_FING 281 303
FT ZN_FING 309 321
FT ZN_FING 337 359
FT ZN_FING 365 387
FT ZN_FING 393 415
FT ZN_FING 421 443
FT ZN_FING 449 471
FT ZN_FING 477 499
FT ZN_FING 505 527
FT ZN_FING 533 555
FT ZN_FING 561 583
FT ZN_FING 589 611
FT CONFLICT 71 71 S -> SGP (in Ref. 2).
FT CONFLICT 85 85 Q -> H (in Ref. 2).
FT CONFLICT 115 115 S -> R (in Ref. 2).
FT CONFLICT 132 132 C -> S (in Ref. 2).
FT CONFLICT 149 149 V -> D (in Ref. 2).
FT NON_TER 1
SQ SEQUENCE 616 AA; 70572 MW; DF309883AB61160A CRC64;

Query Match 72.2%; Score 319; DB 1; Length 616;
Best Local Similarity 70.1%; Pred. No. 1.5e-25;
Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPKYKCEGKAFQSSSLTKHKKIHHTGKPKYKQC 60
Db 451 CEECGKAFNQSSSLTKHKKIHHTGKPKYKCEGKAFNQSSSLTKHKKIHHTGKPKYKCEC 510

QY 61 GKAFGCPNLRHGRTH 77
Db 511 GKAFNQSSSLTKHKKIH 527

RESULT 5
Q9H9X5 HUMAN PRELIMINARY; PRT; 173 AA.
AC Q9H9X5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ12488.
GN Name=ZNF505;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
EX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
```

```
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Iogiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AK022550; BAB14093.1; -; mRNA.
DR HSSP; P08047; 1SP2.
DR Ensembl; ENSG00000081665; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR006025; Pept M Zn BS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2_5.
DR PRINTS; PRO0048; ZINC_FINGER.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 6.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
KW Metal-binding; Nuclear protein; zinc; Zinc-finger.
SQ SEQUENCE 173 AA; 19543 MW; 79332B6AC4C36424 CRC64;

Query Match 71.3%; Score 315; DB 2; Length 173;
Best Local Similarity 70.1%; Pred. No. 1.1e-25;
Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPKYKCEGKAFQSSSLTKHKKIHHTGKPKYKQC 60
Db 92 CEECGKAFNQSSSLTKHKKIHHTGKPKYKCEGKAFNQSSSLTKHKKIHHTGKPKYKCEC 151

QY 61 GKAFGCPNLRHGRTH 77
Db 152 GKAFISPSLSRHEIH 169

RESULT 6
Q59FB0 HUMAN PRELIMINARY; PRT; 403 AA.
AC Q59FB0
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE PREDICTED: KBAB domain only 2 variant (Fragment).
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```

```

RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RA "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AB209550; BAD2787.1; -, mRNA.
KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 403 AA; 46567 MW; 494EC7E4DF3401A0 CRC64;

Query Match 71.3%; Score 315; DB 2; Length 403;
Best Local Similarity 68.8%; Pred. No. 2.6e-25;
Matches 53; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFQKSNLRHQRHTGKPYKCECGKAFQTSNLTGKHKHHTGKPYKCKQC 60
Db 320 CEECGKAFQTSAGLHKHRRHTGKPYKCECGKAVTSSNLTGKHKHHTGKPYKCKEC 379
QY 61 GKAFGCPNLRHGRTH 77
Db 380 GKAFNWSDDLKHKRIH 396

RESULT 7
Q4G180 HUMAN
ID Q4G180_HUMAN PRELIMINARY; PRT; 503 AA.
AC Q4G180;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE ZNF486 protein (Fragment).
GN Name=ZNF486;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner B., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullány S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027608; AAH27608.1; -, mRNA.
FT NON_TER 1
FT NON_TER 503 503
```

```

SQ SEQUENCE 503 AA; 57886 MW; 67B3EF2978354A2A CRC64;

Query Match 71.3%; Score 315; DB 2; Length 503;
Best Local Similarity 68.8%; Pred. No. 3.3e-25;
Matches 53; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFQKSNLRHQRHTGKPYKCECGKAFQTSNLTGKHKHHTGKPYKCKQC 60
Db 420 CEECGKAFQTSAGLHKHRRHTGKPYKCECGKAVTSSNLTGKHKHHTGKPYKCKEC 479
QY 61 GKAFGCPNLRHGRTH 77
Db 480 GKAFNWSDDLKHKRIH 496

RESULT 8
Q8NAG64 HUMAN
ID Q8NAG64_HUMAN PRELIMINARY; PRT; 530 AA.
AC Q8NAG64;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ35804.
GN Name=ZNF568;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori K., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arima M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoigaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AK093123; BAC04064.1; -, mRNA.
DR HSSP; P07248; 1PAA.
DR Ensembl; ENSG00000198453; Homo sapiens.
DR HGNC; HGNC:25392; ZNF568.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

```
DR GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 12.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 13.
FT NON_TER 530 530
SQ SEQUENCE 530 AA; 61275 MW; 9491D3767BF5C5E4 CRC64;

Query Match 71.3%; Score 315; DB 2; Length 530;
Best Local Similarity 71.4%; Pred. No. 3.4e-25;
Matches 55; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 CKDCGKAPIQKSNLIRHQRTHTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 60
Db 412 CSECGKAFQSSNLTGKKIHTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 471
QY 61 GKAFGCPNLRHGRTH 77
Db 472 GKAFSQRQNLLEHEKIH 488

RESULT 9
Q6ZMR0 HUMAN
ID Q6ZMR0_HUMAN PRELIMINARY; PRT; 563 AA.
AC Q6ZMR0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16755.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cerebellum;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Maehuo Y., Nagai K., Isegai T.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131527; BAD18665.1; -; mRNA.
DR HSSP; P08046; 1A1G.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 14.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 14.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.

DR GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 12.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 13.
FT NON_TER 530 530
SQ SEQUENCE 530 AA; 61275 MW; 9491D3767BF5C5E4 CRC64;

Query Match 71.3%; Score 315; DB 2; Length 530;
Best Local Similarity 71.4%; Pred. No. 3.4e-25;
Matches 55; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 CKDCGKAPIQKSNLIRHQRTHTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 60
Db 412 CSECGKAFQSSNLTGKKIHTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 471
QY 61 GKAFGCPNLRHGRTH 77
Db 472 GKAFSQRQNLLEHEKIH 488

RESULT 10
ZNF91 HUMAN
ID ZNF91_HUMAN STANDARD; PRT; 1191 AA.
AC Q05481;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger protein 91 (Zinc finger protein HTF10) (HPF7).
GN Name=ZNF91;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9323677; PubMed=8467795;
RA Bellefroid E.J., Marine J.C., Ried T., Lecocq P.J., Riviere M.,
RA Amemiya C.T., Poncelet D.A., Coulie P.G., de Jong P.J., Szpirer C.,
RA Ward D.C., Martial J.A.;
RT "Clustered organization of homologous KRAB zinc-finger genes with
RT enhanced expression in human T lymphoid cells.";
RL EMBO J. 12:1363-1374(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 15-204.
RX MEDLINE=91219421; PubMed=2023909;
RA Bellefroid E.J., Poncelet D.A., Lecocq P.J., Revelant O.,
RA Martial J.A.;
RT "The evolutionarily conserved Kruppel-associated box domain defines a
RT subfamily of eukaryotic multifingered proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3608-3612(1991).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC family.
CC -!- SIMILARITY: Contains 36 C2H2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC -!- CAUTION: The sequence from position 1159 to the C-terminal is
CC derived from the translation of an Alu repeat.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L11672; AAA59469.1; -; mRNA.
CC EMBL; M61871; AAA58672.1; ALT_SEQ; mRNA.
CC PIR; S35305; S35305.
CC HSSP; P08047; 1SP2.
CC SMR; Q05481; 291-372.
CC Ensembl; ENSG00000167232; Homo sapiens.
CC HGNC; HGNC:13166; ZNF91.
CC MIM; 603971; -.
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003700; F:transcription factor activity; NAS.
CC GO; GO:0008270; F:zinc ion binding; NAS.
```


10-MAY-2005 (Rel. 47, Last annotation update)
Zinc finger protein OZF.
Name=ZNF146; Synonyms=OZF;
Homo sapiens (Human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
Homo.
NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Mammary gland;
RC MEDLINE=94149744; PubMed=8107129;
RA le Chalony C., Prosperi M.-T., Haluza R., Apiou F., Dutrillaux B.,
RA Goubin G.;
RT "The OZF gene encodes a protein consisting essentially of zinc finger
RT motifs.";
RL J. Mol. Biol. 236:399-404(1994).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=21203548; PubMed=11306801;
RA Pibouin L., Villaudy J., Prosperi M., Goubin G.;
RT "Genomic organization and promoter identification of ZNF146, a gene
RT encoding a protein consisting solely of zinc finger domains.";
RL Cytogenet. Cell Genet. 92:80-84(2001).
RN [3]
RN CHARACTERIZATION.
RP MEDLINE=96270754; PubMed=8665923;
RA Ferbus D., le Chalony C., Prosperi M.-T., Muleris M.,
RA Vincent-Salomon A., Goubin G.;
RT "Identification, nuclear localization, and binding activities of OZF,
RT a human protein solely composed of zinc-finger motifs.";
RL Eur. J. Biochem. 236:991-995(1996).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Liver, skeletal and heart muscle, mammary
CC cells. Very low levels in brain, lung, placenta and kidney.
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC family.
CC -!- SIMILARITY: Contains 10 C2H2-type zinc fingers.
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/OZFID267.html"
CC
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; J03194; CAA49844.1; -; mRNA.
DR EMBL; AJ011806; CAB41967.1; -; Genomic_DNA.
DR PIR; S43826; S43826.
DR HSSP; P08047; 1SP2.
DR TRANSFAC; T02323; -.
DR Ensembl; ENSG00000167635; Homo sapiens.
DR HGNC; HGNC:12931; ZNF146.
DR MIM; 601505; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003677; F:DNA binding; TAS.
DR GO; GO:0008201; F:heparin binding; TAS.
DR GO; GO:0008270; F:zinc ion binding; TAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 10.
DR PRINTS; PR00048; ZINC_FINGER.
DR PRODOM; PD000003; Znf_C2H2; 10.
DR SMART; SM00355; Znf_C2H2; 10.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 10.
KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Zinc;
KW Zinc-finger.
FT ZN_FING 16 38 C2H2-type 1.
FT ZN_FING 44 66 C2H2-type 2.
FT ZN_FING 72 94 C2H2-type 3.

FT ZN_FING 100 122 C2H2-type 4.
FT ZN_FING 128 150 C2H2-type 5.
FT ZN_FING 156 178 C2H2-type 6.
FT ZN_FING 184 206 C2H2-type 7.
FT ZN_FING 212 234 C2H2-type 8.
FT ZN_FING 240 262 C2H2-type 9.
FT ZN_FING 268 290 C2H2-type 10.
SQ SEQUENCE 292 AA; 33280 MW; DD9B7BA8D43F0C9C CRC64;

Query Match 71.0%; Score 314; DB 1; Length 292;
Best Local Similarity 70.1%; Pred. No. 2.4e-25;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 60
DB 102 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 161

QY 61 GKAFGCPNLRHGRTH 77
DB 162 GTAFGQKKYLIKQNIH 178

RESULT 12
Q5RAA6_PONPY PRELIMINARY; PRT; 292 AA.
AC Q5RAA6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459E0928.
GN Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR859112; CAH91304.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR007086; Znf_C2H2.
DR Pfam; PF00096; zf_C2H2_10.
DR PRINTS; PR000003; Znf_C2H2; 10.
DR SMART; SM00355; Znf_C2H2; 10.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 10.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ SEQUENCE 292 AA; 33309 MW; 8F0626BAB261BAFC CRC64;

Query Match 71.0%; Score 314; DB 2; Length 292;
Best Local Similarity 70.1%; Pred. No. 2.4e-25;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 60
DB 102 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 161

QY 61 GKAFGCPNLRHGRTH 77
DB 162 GTAFGQKKYLIKQNIH 178

RESULT 13
Q5RFP4_PONPY PRELIMINARY; PRT; 292 AA.
AC Q5RFP4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469K1922.
GN Name=DKFZp469K1922;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG The German cDNA Consortium;
RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR857109; CAH89413.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf_C2H2_10.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 10.
DR SMART; SM00355; Znf_C2H2; 10.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 10.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ SEQUENCE 292 AA; 33308 MW; D5062375D261B971 CRC64;

Query Match 71.0%; Score 314; DB 2; Length 292;
Best Local Similarity 70.1%; Pred. No. 2.4e-25;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 60
DB 102 CKDCGKAFIQSNLIRHQRTHTGKPYKCEGKAFQTSNLTGHKTKHTGKPYKCKQC 161

QY 61 GKAFGCPNLRHGRTH 77
DB 162 GTAFGQKKYLIKQNIH 178

RESULT 14
ZN502_HUMAN
ID ZN502_HUMAN STANDARD; PRT; 544 AA.
AC Q8TBZ5;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger protein 502.
GN Name=ZNF502;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins C.M., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carininci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA	Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.; and
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC	-1- FUNCTION: May be involved in transcriptional regulation.
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	-1- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC	family.
CC	-1- SIMILARITY: Contains 14 C2H2-type zinc fingers.
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	EMBL; BC028377; AAH28377.1; -; mRNA.
DR	HSSP; P08047; 1SP2.
DR	SNR; Q8TB25; 348-429.
DR	Ensembl; ENSG00000196653; Homo sapiens.
DR	HGN; HGNC:33718; ZNF502.
DR	InterPro; IPR007087; Znf_C2H2.
DR	InterPro; IPR007086; Znf_C2H2_sub.
DR	Pfam; PF00096; zf-C2H2; 14.
DR	PRINTS; PR00048; ZINCFINGER.
DR	ProDom; PD000003; Znf_C2H2; 14.
DR	SMART; SM00355; Znf_C2H2; 14.
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.
DR	PROSITE; PS00157; ZINC_FINGER_C2H2_2; 14.
KW	DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
KW	Transcription regulation; Zinc; Zinc-finger.
FT	ZN_FING 155 177 C2H2-type 1.
FT	ZN_FING 183 205 C2H2-type 2.
FT	ZN_FING 211 233 C2H2-type 3.
FT	ZN_FING 239 261 C2H2-type 4.
FT	ZN_FING 267 289 C2H2-type 5.
FT	ZN_FING 295 317 C2H2-type 6.
FT	ZN_FING 323 345 C2H2-type 7.
FT	ZN_FING 351 373 C2H2-type 8.
FT	ZN_FING 379 401 C2H2-type 9.
FT	ZN_FING 407 429 C2H2-type 10.
FT	ZN_FING 435 457 C2H2-type 11.
FT	ZN_FING 463 485 C2H2-type 12.
FT	ZN_FING 491 513 C2H2-type 13.
FT	ZN_FING 519 541 C2H2-type 14.
SQ	SEQUENCE 544 AA; 63920 MW; 63437B670A6B22D0 CRC64;

```

Query Match      71.0%; Score 314; DB 1; Length 544;
Best Local Similarity 70.1%; Pred. No. 4.5e-25;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKQCGKAFIQKSNLIRHQRTHTGKPYKCECGKAPTQSSNLTKHKXHTGKEPKYKQC 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 CSEGGKAFIQSICLIRHQRSHTGKPYKCEGKFNQNTCLTQHMRIHTGKEPKYKCEC 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Qy	61 GKAFGCPNLRHRHRTTH	77
	:	
Dd	469 GKFAHSSSLTEHRHTH	485

RESULT 15

Q6AY11_RAT PRELIMINARY; PRT; 589 AA.
ID Q6AY11_RAT PRELIMINARY; PRT; 589 AA.
AC Q6AY11;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zinc finger protein 386 (Kruppel-like).
GN Name=Znf386;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=22388457; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RL Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC078841; AAH78841.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf_C2H2; 9.
DR PRINTS; PR00048; ZINC_FINGER.
DR ProDom; PD000003; Znf_C2H2; 9.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 9.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 9.
SQ SEQUENCE 589 AA; 68356 MW; 3B070638297C903 CRC64;

```
Query Match          71.0%; Score 314; DB 2; Length 589;  
Best Local Similarity   66.2%; Pred. No. 4.9e-35;  
Matches      51; Conservative    12; Mismatches     14; Indels       0; Gaps        0;
```

	Qy	1	CXDCGKAFQTQSNIIBHORTHTECKPYKEECCGAFTOSSNLTKHKKIHTGEKPKYCKQC	60
Db	456	CABCGRGSFQNSTLSQHRIHTGEKPYKECQGCAFTQCSSLRKHORIHTGEKPKYCCEB	515	
Qy	61	GKAFCGPCSNLRRHGRTH	77	

Db 516 GRAFNCRSSFTKKRIH 532

Search completed: January 9, 2006, 20:08:48
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 11, 2006, 20:01:04 ; Search time 3642 Seconds
(without alignments)
1201.797 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442
Sequence: 1 CKDCGKAFIQKSNLIRHQT.....KQCGKAFGCPNLRHRGRTH 77

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144453_20373/app_query.fasta_1.263
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pt0 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10669861 @CEN 1 1 4939 @runat 09012006_144453_20373 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	329	74.4	2562	6	CQ728633 Sequence
2	329	74.4	158404	8	AC092835 Homo sapi
3	329	74.4	294915	14	AC155005 Bos tauri

SUMMARIES

4	327	74.0	672	6	AX721061	AX721061 Sequence
5	327	74.0	2192	6	AX540402	AX540402 Sequence
6	327	74.0	2275	6	AX468105	AX468105 Sequence
7	327	74.0	2280	6	AX835301	AX835301 Sequence
8	327	74.0	2280	8	AC098300	AC098300 Homo sapi
9	327	74.0	2323	8	BC047412	BC047412 Homo sapi
10	327	74.0	178184	8	AC008770	AC008770 Homo sapi
11	327	74.0	179581	14	AC009397	AC009397 Homo sapi
12	327	74.0	180510	8	AC022415	AC022415 Homo sapi
13	325	73.5	3300	6	AX202430	AX202430 Sequence
14	322	72.9	216979	14	AC103247	AC103247 Rattus no
15	322	72.9	238307	14	AC133610	AC133610 Rattus no
16	321	72.6	330	6	AX202476	AX202476 Sequence
17	321	72.6	3300	6	AX202426	AX202426 Sequence
18	321	72.6	3300	6	AX202427	AX202427 Sequence
19	319	72.2	2667	6	AX834102	AX834102 Sequence
20	319	72.2	2667	8	AC096342	AC096342 Homo sapi
21	319	72.2	109478	8	AC007204	AC007204 Homo sapi
22	318	71.9	197553	14	AC146103	AC146103 Pan trogl
23	316	71.5	2253	6	CQ779570	CQ779570 Sequence
24	316	71.5	127811	14	AC008375	AC008375 Homo sapi
25	316	71.5	138491	9	AL929042	AL929042 Mouse DNA
26	316	71.5	146877	14	AC026487	AC026487 Homo sapi
27	316	71.5	169602	14	AC079567	AC079567 Mus muscu
28	316	71.5	202768	14	AC092345	AC092345 Homo sapi
29	316	71.5	215962	8	AC010615	AC010615 Homo sapi
30	316	71.5	239441	14	AC128484	AC128484 Rattus no
31	316	71.3	827	6	BD147256	BD147256 Primer fo
32	315	71.3	827	6	AX867194	AX867194 Sequence
33	315	71.3	882	6	CQ734128	CQ734128 Sequence
34	315	71.3	1449	6	BD157779	BD157779 Primer fo
35	315	71.3	1449	6	AX879332	AX879332 Sequence
36	315	71.3	1449	8	AC022550	AC022550 Homo sapi
37	315	71.3	1512	8	BC027608	BC027608 Homo sapi
38	315	71.3	1914	8	AX747939	AX747939 Sequence
39	315	71.3	1914	8	AK093123	AK093123 Homo sapi
40	315	71.3	1914	8	AK093123	AK093123 Homo sapi
41	315	71.3	1997	6	CQ726025	CQ726025 Sequence
42	315	71.3	2648	8	HSN807672	HSN807672 Homo sapi
43	315	71.3	2664	6	AX833387	AX833387 Sequence
44	315	71.3	2664	8	AK095062	AK095062 Homo sapi
45	315	71.3	3096	6	CQ731632	CQ731632 Sequence

ALIGNMENTS

RESULT 1	CQ728633	Sequence 14567 from Patent WO02068579.	2562 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ728633	Sequence 14567 from Patent WO02068579.				
DEFINITION	CQ728633					
ACCESSION	CQ728633.1	GI:42297975				
VERSION	CQ728633.1	GI:42297975				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
AUTHORS	Kites, such as nucleic acid arrays, comprising a majority of					
TITLE	humanexons or transcripts, for detecting expression and other uses					
JOURNAL	thereof					
FEATURES	PE Corporation (NY) (US)					
source	Location/Qualifiers					
	1..2562					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

ORIGIN

Alignment Scores:

```

Pred. No.: 1.89e-27 Length: 2562
Score: 329.00 Matches: 56
Percent Similarity: 85.71% Conservative: 10
Best Local Similarity: 72.73% Mismatches: 11
Query Match: 74.43% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x CQ728633 (1-2562)

QY 1 CysLysAspCysGlyLysAlaPheHleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 2248 TGCAATGATTGTCGCAAGACCTTCAGTCACCGCTCAGCCCTATTTCGTCATCAGAGAACA 2307

QY 21 HisThrGlyLysLysProTyrLysCysGluGlnCysGlyLysAlaPheThrGlnSerSer 40
Db 2308 CACACTGGAGAGAACCTTACAAATGTATAGATTGTGGAAGAGCAATTCAGCCAGAGCTCA 2367

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2368 TCTCTTCAAGACATCAGAAAACCTCACACTGGAGAAAAGCCCTACAAAGTGAAGGAATGT 2427

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2428 GGAAGAGCCTTAGCCAGAGTTTCCTCCTCTCAACATCAGAAAACCTCAT 2478

RESULT 2
AC092835 LOCUS AC092835 158404 bp DNA linear PRI 15-APR-2005
DEFINITION Homo sapiens BAC clone RP11-11P22 from 2, complete sequence.
ACCESSION AC092835
VERSION AC092835.4 GI:19071676
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 158404)
AUTHORS Doebber,A., Nguyen,C. and Haglund,K.
TITLE The sequence of Homo sapiens BAC clone RP11-11P22
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 158404)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 158404)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 158404)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 158404)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Mar 2, 2002 this sequence version replaced gi:18250106.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0011P22
-----

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-708D7; the clone sequenced to the right is RP11-468G5, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-11P22.

There is a transposon that has been omitted from the submitted sequence. The transposons should insert after base position 141089. There is a PCR only region from 145576 to 145630.

Polymorphisms have been identified between AC093621 and AC092835. Data from AC093621, AC103563 and AC026626 was used to finish this clone, AC092835.

FEATURES

```

location/Qualifiers
1. 158404
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-11P22"
/clone_lib="RPCI-11"
/complement(5713..40509)
/gene="MRPS5"
complement(join(5713..6086,8891..9027,18979..19041,
19342..19399,20182..20228,23145..23235,24928..24962,
26680..26913,28421..28546,33571..33708,36370..36450,
40239..40509))
/gene="MRPS5"
complement(join(5862..6086,8891..9027,18979..19041,
19342..19399,20182..20228,23145..23235,24928..24962,
26680..26913,28421..28546,33571..33708,36370..36450,
40239..40296))
/gene="MRPS5"
/note="Homo sapiens mitochondrial ribosomal protein S5
(MRPS5), nuclear gene encoding mitochondrial protein,
mRNA; H_NH0011P22.1
This gene was based on gi(16554614)"
/codon_start=1
/product="unknown"
/db_xref="GI:62630235"
/translation="MATAVRAGCLPVLCSGTAGHLGROCSLNTLPAASILAKWSVL
GNHLSIGTRDTHPYASLSALQTCQISPSHLMQOQRPYSEFTKLTADLWKGA
LAETGAGAKGRKRTKKKKKDLNRQIIGEGRYFLMPLGNVFLMKNQAVQITQR
SKEQEKVADMTIQREBWRKKMKVGRGWSNGSGISGLPDPGPGCQYEDF
DTRIILEVNVFTTAKGRKKSIRLVAVNGKGAAGFSIGKATDRMDAFKAKRV
HHLHYIERYEDTIFHDISLRFKRTHIMKKQPKGYGLRCHRAITTCRLIGKMDYA
KVGSGINWLSLTQGLFRLSRQETHQQLADKGLHVVIREECGLPLTIWVASPRGLR

```

KDPEDPEVDVVDKLDHEDVKTACQMKRSVWSNLRKAAT"
 8404..8639
 /notes="CpG island (%GC=61.4, o/e=0.95, #CpGs=20)"
 39850..40711
 /notes="CpG island (%GC=69.5, o/e=0.85, #CpGs=92)"
 complement(66160..78023)
 /gene="FLJ14457"
 complement(join(66160..68772,71179..71274,71638..71764,
 75856..75944,77745..78023))
 /gene="FLJ14457"
 complement(join(67787..68772,71179..71274,71638..71758))
 /gene="FLJ14457"
 /note="Homo sapiens zinc finger protein 514 (ZNF514),
 mRNA.; H.NH0011P22.2
 This gene was based on gi(14249455)"
 /codon_start=1
 /product="unknown"
 /protein_id="AAK8981.1"
 /db_xrefs="GI:62630236"
 /translation="MTFEDVAFVSQWNGQLNPAQKDLRYEVMLENFRNLAILGLLV
 SKPYVLCQEEGEPFVREISTGAHDKRKSKEKSMPSGWIISKBELFQVVSVEK
 HIQDLVQLKAAACDQGLEMQOIKQERHLKQKSTIHKSATLISRDYKMGFGFSL
 GLRVLVNOHSILMGSGYKCDTPEPQTGLGNNSTHPEKSKCKNECKGSPHPOSE
 LRHORCHTGEKPYECSDCGRAFGLHSLIKQRTHTGKPYECSCGGRAPQSSSLV
 LHYRFTGKPYKNCBGRAFGLHSLIKQRTHTGKPYECRECGRTFSQSSSLVH
 YRPHTEGKPYKNCBGRAFGLHSLIKQRTHTGKPYECRECGRTFSQSSSLVH
 SHAGKKTLL"
 77563..77841
 /notes="CpG island (%GC=76.1, o/e=0.79, #CpGs=108)"
 83918..84593
 /notes="CpG island (%GC=61.5, o/e=0.98, #CpGs=57)"
 83950..101481
 /gene="ZNF2"
 join(83950..84365,94693..94764,95988..96114,98664..98777,
 99608..101481)
 /gene="ZNF2"
 99733..100614
 /gene="ZNF2"
 /note="Homo sapiens zinc finger protein 2 (ZNF2),
 mRNA.; H.NH0011P22.3
 This gene was based on gi(20304090)"
 /codon_start=1
 /product="unknown"
 /protein_id="AAK8982.1"
 /db_xrefs="GI:62630237"
 /translation="MTEKQSPGSETRKSLRDKGLRPRRSALGREILTKERHOECS
 DCGKTPFDHSLTRHRTHTGKPYDCRCGKAFSRSLRSLHLSHTGESPYECVC
 SKAFDRSLTRHRTHTGKPYDCRCGKAFSRSLRSLHLSHTGESPYECVC
 AFSQSLTRHRTHTGKPYDCRCGKAFYGVSLNRHOKAHAGDPYQCNCEKAF
 FDRSLTRHRTHTGKPYDCRCGKAFYGVSLNRHOKAHAGDPYQCNCEKAF
 KSVIQRHRYAKQID"
 102198..102424
 /notes="CpG island (%GC=58.6, o/e=0.94, #CpGs=17)"
 125652..126308
 /notes="CpG island (%GC=70.6, o/e=0.87, #CpGs=68)"
 151899..152168
 /notes="CpG island (%GC=60.7, o/e=0.74, #CpGs=20)"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.9e-25 Length: 158404
 Score: 329.00 Matches: 56
 Percent Similarity: 85.71% Conservative: 10
 Best Local Similarity: 72.73% Mismatches: 11
 Query Match: 74.43% Indels: 0
 DB: 8 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x AC092835 (1-158404)
 QY 1 CysLysAspCysGlyLysAlaPheLleGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 DB 133797 TGCATGATGTTGGCCAAAGCCCTTACGTGACCGCTCAGCCCTTATGTCATCAGAGAAC 133856

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
 DB 133857 CACACTGGAGAGAACTTACAAATGTAAGATTGTGGAAGCATTACGCCAGACTCA 133916
 QY 41 AsnLeuThrLysHisLysLysLeHisThrGlyLysProTyrLysCysGlnCys 60
 DB 133917 TCTCTTACAAAGCATCAGAAACTCACCCTGGAGAAAGCCCTACAGTGTAGGAATGT 133976
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgGlyHisGlyArgThrHis 77
 DB 133977 GGAAAGCCCTTACGAGAGTTCATCCCTCTCTCAACATCAGAAAAAATCAT 134027
 RESULT 3
 AC155005
 LOCUS
 DEFINITION Bos taurus clone CH240-36J15, *** SEQUENCING IN PROGRESS ***, 48
 unordered pieces.
 AC155005
 VERSION AC155005.3 GI:68265223
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 Muzny,D,Marie., Metzker,M,Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Eacotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuewa,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
 Nwakoelamch,O., Okwuonu,G., Olarpunsaagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,
 Plopper,F., Poinexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanai,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von

Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 294915)
 Worley,K.C.
 Direct Submission
 Submitted (07-JAN-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 294915)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUN-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 28, 2005 this sequence version replaced gi:57334887.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FCCE
 Center clone name: CH240-36J15
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 262862 bases at least Q40
 Consensus quality: 266947 bases at least Q30
 Consensus quality: 270751 bases at least Q20
 Estimated insert size: 269122; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one 'clone'
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 48 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2521: contig of 2521 bp in length
 * 2522 2571: gap of 50 bp
 * 2572 5531: contig of 2960 bp in length
 * 5532 5631: gap of unknown length
 * 5632 7766: contig of 2135 bp in length
 * 7767 7816: gap of 50 bp
 * 7817 15053: contig of 7237 bp in length
 * 15054 15103: gap of 50 bp
 * 15104 19441: contig of 4338 bp in length
 * 19442 19491: gap of 50 bp
 * 19492 23159: contig of 3668 bp in length
 * 23160 23259: gap of unknown length
 * 23260 24997: contig of 1738 bp in length
 * 24998 25047: gap of 50 bp
 * 25048 45806: contig of 20759 bp in length
 * 45807 45856: gap of 50 bp
 * 45857 53466: contig of 7610 bp in length

* 53467 53516: gap of 50 bp
 * 53517 58591: contig of 5075 bp in length
 * 58592 58641: gap of 50 bp
 * 58642 66086: contig of 7445 bp in length
 * 66087 66136: gap of 50 bp
 * 66137 81849: contig of 15713 bp in length
 * 81850 81899: gap of 50 bp
 * 81900 84423: contig of 2524 bp in length
 * 84424 96159: gap of 11736 bp
 * 96160 100621: contig of 4462 bp in length
 * 100622 100671: gap of 50 bp
 * 100672 104944: contig of 4273 bp in length
 * 104945 104994: gap of 50 bp
 * 104995 112201: contig of 7207 bp in length
 * 112202 112251: gap of 50 bp
 * 112252 132363: contig of 20111 bp in length
 * 132363 132412: gap of 50 bp
 * 132413 140044: contig of 7632 bp in length
 * 140045 140094: gap of 50 bp
 * 140095 162586: contig of 22492 bp in length
 * 162587 162636: gap of 50 bp
 * 162637 170570: contig of 7934 bp in length
 * 170571 171490: gap of 920 bp
 * 171491 175896: contig of 4406 bp in length
 * 175897 175946: gap of 50 bp
 * 175947 183357: contig of 5411 bp in length
 * 183358 183373: gap of 1016 bp
 * 183374 183788: contig of 1415 bp in length
 * 183789 183988: gap of 200 bp
 * 183989 212516: contig of 28528 bp in length
 * 212517 214190: gap of 1674 bp
 * 214191 216319: contig of 2129 bp in length
 * 216320 216419: gap of unknown length
 * 216420 222050: contig of 5631 bp in length
 * 222051 222100: gap of 50 bp
 * 222101 226487: contig of 4387 bp in length
 * 226488 226587: gap of unknown length
 * 226588 228131: contig of 1544 bp in length
 * 228132 228365: gap of 234 bp
 * 228366 231899: contig of 3534 bp in length
 * 231900 231999: gap of unknown length
 * 232000 233871: contig of 1872 bp in length
 * 233872 233921: gap of 50 bp
 * 233922 240058: contig of 6137 bp in length
 * 240059 240108: gap of 50 bp
 * 240109 254341: contig of 14233 bp in length
 * 254342 254441: gap of unknown length
 * 254442 259967: contig of 5526 bp in length
 * 259968 260017: gap of 50 bp
 * 260018 263868: contig of 3851 bp in length
 * 263869 263968: gap of unknown length
 * 263969 265021: contig of 1053 bp in length
 * 265022 265121: gap of unknown length
 * 265122 266132: contig of 1011 bp in length
 * 266133 266232: gap of unknown length
 * 266233 267543: contig of 1311 bp in length
 * 267544 267643: gap of unknown length
 * 267644 268658: contig of 1015 bp in length
 * 268659 268758: gap of unknown length
 * 268759 270271: contig of 1513 bp in length
 * 270272 270371: gap of unknown length
 * 270372 271638: contig of 1267 bp in length
 * 271639 271738: gap of unknown length
 * 271739 273324: contig of 1586 bp in length
 * 273325 273424: gap of unknown length
 * 273425 274728: contig of 1304 bp in length
 * 274729 274828: gap of unknown length
 * 274829 276302: contig of 1474 bp in length

Alignment Scores: 3.82e-25 Length: 294915
 Pred. No.: 329.00 Matches: 56
 Score: 85.71% Conservative: 10
 Percent Similarity:

Best Local Similarity: 72.73% Mismatches: 11
 Query Match: 74.43% Indels: 0
 DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC155005 (1-294915)

QY 1 CysLysAspCysGlyLysAlaPheGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 Db 44564 TGAACGATTGTGCCAAAGCCTTCAGTGACCGTTCAGCCCTTATCGCTCATCAGAGAAC 44623
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
 Db 44624 CACACTGGAGAGAAACCTTACAAAGTGAAGAGTGGAAAGCCTTCAGCCAGAGCTCA 44683
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 44684 TCCTTACAAGCATCAGAAACTCAGCTGGAGAAACCCCTATTAAGTGAAGGAATGT 44743
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
 Db 44744 GGAAAGCCTTCAGCCAGAGTTCATCCCTTCTCAACATCAGAAACTCAT 44794

RESULT 4
 LOCUS AX721061 672 bp mRNA linear PAT 07-MAY-2003
 DEFINITION Sequence 21 from Patent WO0220754.
 ACCESSION AX721061
 VERSION AX721061.1 GI:30421897
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Lincoln, S.E., Altus, C.M., Dufour, G.E., Chalup, M.S., Hillman, J.L., Jones, A., Yu, J.Y., Wright, R.J., Gietzen, D., Liu, Toomy, F., Yap, P., Dahl, C.R., Momiya, M., Bradley, D., Rohatgi, S., Harris, B., Roseberry, A.M., Gerstin, E.H., Peralta, C.H., David, M., Panzer, S., Flores, V., Daffo, A., Marwaha, R., Chen, A., Chang, S.C. and Inman, R.R.
 TITLE Molecules for diagnostics and therapeutics
 JOURNAL Patent: WO 0220754-A 21 14-MAR-2002;
 Incyte Genomics, Inc. (US)

FEATURES
 source
 1..672
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: LG:1080545.1:2000SEP08"

ORIGIN

Alignment Scores:
 Pred. No.: 7.15e-28 Length: 672
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX721061 (1-672)

QY 1 CysLysAspCysGlyLysAlaPheGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 Db 311 TGAACGATTGTGCCAAAGCCTTCAGTGACCGTTCAGCCCTTATCGCTCATCAGAGAAC 370
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
 Db 371 CACACTGGAGAGAAACCTTACAAAGTGAAGAGTGGAAAGCCTTCAGCCAGAGCTCA 430
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 431 CGACTTCAAAATGATGGAAGGAGTTCACACTGGAGAGAAACCCGTATAAATGTAAGCAATGT 490

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
 Db 491 GGGAAAGCCTTTGGATGTCCCTCAAACTTCGAGGCATGGAAGGACTCAC 541

RESULT 5
 LOCUS AX540402 2192 bp DNA linear PAT 23-NOV-2002
 DEFINITION Sequence 14 from Patent WO02055738.
 ACCESSION AX540402
 VERSION AX540402.1 GI:25273419
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Panzer, S.R., Lincoln, S.E., Altus, C.M., Dufour, G.E., Hillman, J.L., Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Daffo, A., Marwaha, R., Chen, A.J., Chang, S.C., Gerstin, E.H., Peralta, C.H., David, M.H., and Lewis, S.A.
 TITLE Molecules for disease detection and treatment
 JOURNAL Patent: WO 02055738-A 14 18-JUL-2002;
 Incyte Genomics, Inc. (US)

FEATURES
 source
 1..2192
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: LI:1175131.1:2001JAN12"

ORIGIN

Alignment Scores:
 Pred. No.: 2.68e-27 Length: 2192
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX540402 (1-2192)

QY 1 CysLysAspCysGlyLysAlaPheGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 Db 1845 TGAACGATTGTGCCAAAGCCTTCAGTGACCGTTCAGCCCTTCAATGATGGAAGACT 1904
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
 Db 1905 CACACTGGAGAGAAACCTTATGAATGTAAGCAGTGTGGGAAGTCTTTGGATGCTCG 1964
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 1965 CGACTTCAAAATGATGGAAGGACTCACACTGGAGAGAAACCCGTATAAATGTAAGCAATGT 2024
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
 Db 2025 GGGAAAGCCTTTGGATGTCCCTCAAACTTCGAGGCATGGAAGGACTCAC 2075

RESULT 6
 LOCUS AX468105 2275 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 28 from Patent WO0250279.
 ACCESSION AX468105
 VERSION AX468105.1 GI:21900979
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Baughin, M.R., Lu, Y., Arvizu, C., Ramkumar, J., Yao, M.G., Policky, J.L., Wallia, N.K., Tribouley, K.M., Yue, H., Batra, S.,

Ding,L., Lal,P.G., Borowsky,M.L., Lu,D.A., Gandhi,A.R.,
Griffin,J.A., Xu,Y., Azimzai,Y., Gietzen,K.J., Tang,Y.T.,
Warren,B.A., Mason,P.M., Burford,N., Hafalia,A.J., Lee,E.A.,
Yang,J., Gorvad,A.E., Emerling,B.M., Marquis,J.P., Lee,S.Y.,
Swarnakar,A. and Reddy,R.
Nucleic acid-associated proteins
Patent: WO 0250279-A 28 27-JUN-2002;
Incyte Genomics, Inc. (US)

TITLE
JOURNAL

FEATURES

source
1..2275
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 5435937CB1"

ORIGIN

Alignment Scores:

Pred. No.: 2.8e-27 Length: 2275
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX468105 (1-2275)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
DB 1914 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAATGCATGGAAGGACT 1973
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1974 CACACTGGAGAGAACCTTAAATGTAAGCAGTGTGGAGAGCTTTTGGATGTCCTCG 2033
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 2034 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAACCGTATAAATGTAAGCAATGT 2093
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 2094 GGGAAAGCTTTTGGATGTCTCAAACTTCGAAGCATGGAAGGACTCAC 2144

RESULT 7

AX835301
LOCUS AX835301 2280 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2425 from Patent EP1347046.
ACCESSION AX835301
VERSION AX835301.1 GI:39921436

KEYWORDS

source Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.
Full-length cDNA sequences
Patent: EP 1347046-A 2425 24-SEP-2003;
Research Association for Biotechnology (JP)

TITLE

JOURNAL EP 1347046-A 2425 24-SEP-2003;

FEATURES

source
1..2280
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2.8e-27 Length: 2280
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7

Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX835301 (1-2280)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
DB 1861 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAATGCATGGAAGGACT 1920
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1921 CACACTGGAGAGAACCTTAAATGTAAGCAGTGTGGAGAGCTTTTGGATGTCCTCG 1980
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 1981 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAACCGTATAAATGTAAGCAATGT 2040
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 2041 GGGAAAGCTTTTGGATGTCTCAAACTTCGAAGCATGGAAGGACTCAC 2091

RESULT 8

AX098300
LOCUS AX098300 2280 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40981 fig, clone UTERU2014548, moderately
similar to Homo sapiens ZKI mRNA for Kruppel-type zinc finger
protein.

ACCESSION

VERSION AK098300.1 GI:21758285

KEYWORDS

oligo capping; fig (full insert sequence).

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2280)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

TITLE

JOURNAL

REFERENCE

2 (bases 1 to 2280)

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES
source
1..2280
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2014548"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"
172..2193
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC05279.1"

CDS

CONSRM TITLE	JOURNAL PUBMED REFERENCE AUTHORS	CONSRM TITLE	JOURNAL	REMARK COMMENT
-----------------	---	-----------------	---------	-------------------

Alignment Scores:	
Pred. No.:	2.8e-27
Score:	327.00
Percent Similarity:	81.02%
Best Local Similarity:	72.73%
Query Match:	73.98%
DB:	8
Length:	2380
Matches:	56
Conservative:	7
Mismatches:	14
Indels:	0
Gaps:	0

Qy	1	Cys	Lys	Asp	Cys	Gly	Lys	Ala	Phe	Ile	Gln	Lys	Ser	Asn	Leu	Ile	Arg	His	Gln	Arg	Thr	20
Db	1861	TGT	AAG	CAG	TGT	GGG	AAG	CCG	TTT	TGG	ATC	TGC	CTC	GCA	CTT	CAA	ATG	CAT	TGG	AAG	GAC	1920

QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 |||||
 Db 1921 CACACTGGAGAGAAACCTATGAATGTAAACAGGTGTGGGAAGTCTTTTGTGATGTGCCTCG 1980
 |||||

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyCylulysProTrlYsCysLvsGlnCys 60
 ||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 1981 CGACTTCAAAATGCATGGAGAGACTTCACCTGGAGAGAACCCTATTAANTGTATAGCAATGT 2040

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
|||
Db 2041 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAGGCATCGAAGGACTCAC 2091

RESULT 9	
BC047412	
LOCUS	2323 bp mRNA linear PRI 21-JUL-2005
DEFINITION	Homo sapiens zinc finger protein 433, mRNA (cDNA clone MGC:48500 IMAGE:5295132), complete cds.

ACCESSION BC047412
VERSION BC047412.1 GI:28704058
KEYWORDS MGC.

SOURCE	ORGANISM
Homo sapiens (human)	Homo sapiens
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2323)

REFERENCE
AUTHORS

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zoberg,B., Buetcow,K.H., Schaefcr,C.P., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.S., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Ketterman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Alignment Scores:

Pred. No.:	2.86e-27	Length:	2323
Score:	327.00	Matches:	56
Percent Similarity:	81.82%	Conservative:	7
Best Local Similarity:	72.73%	Mismatches:	14
Query Match:	73.98%	Indels:	0
DB:	8	Gaps:	0

FEATURES
source

```

source
1. .3223
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:48500 IMAGE:5295"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
1. .3223
gene

```

gene

5

translation="MQETPFNMAISGKKWKQNIYVEYNLRNRIIVGERLFPESKEG
HQGBELLQVDPMLKTTTGVGSCSSVYGEVSAHSGNKHIIIDTGHKAYEYQY
GQPKYKICPPNCLSSVQWTHERRASGRXYCEECKTPIISHNQOXRIMURGD
GYKCEKGFALMFLYLIHKRTHTGKPKYQCKQKAFSSHSLRHTHTGKPKY
YKNECGKAFHSSTCLAHKRTHTGKEPKYQCKQKAFSSHSFOHRTHTGKPKY
GKCGKAFKPSVYRHTHRSKPKYCEKCGKLYSLYTSFQNLGHMTGRISSHCK
ICGKAFYSPSLOTHKTHTGKPKYKQCKGAFNSSSFYRHTHTGKPKYCECKQ
GKAFSASLLQTHRTHTGKPYACKGCKPFSNFPFIHRTHTGKPKYCECKGK
TFSUPLFHRHTHTGKTYCKQCKGAFNCSFNSQFVHGRTHTGKPKYCECKGKAF
RNASQLQHRHTHTGKPKYCECKQKAFSGASHLQMGRTHTGKPKYCECKGKFG
ASRLQMGRTHTGKPKYCKQKAFGCPSNLRHRHRTHTGKPKYKQCKGKVFRCSS
OLOVHGHAHCIDTP"

US-10-669-861-2_COPY_33_109 (1-77) x BC047412 (1-2323)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
 Db 1865 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCGCTCACACCTTCAAAATGATGGAAGGACT 1924
 Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 1925 CACACTGGAGAGAAACCCATGAATGTAAGCAGTGTGGAGTCTTTGGATGTGCTCG 1984
 Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyLysGlnCys 60
 Db 1985 CGACTTCAAAATGATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 2044

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77

Db 2045 GGGAAAGCTTTGGATGTCCCTCAAACTTCGAAGCATGGAAGGACTCAC 2095

RESULT 10
 AC008770/c 178184 bp DNA linear PRI 22-MAR-2003
 LOCUS Homo sapiens chromosome 19 clone CTD-2006C1, complete sequence.
 AC008770
 AC008770.7 GI:15887253
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 178184)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 178184)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 178184)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 4 (bases 1 to 178184)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 5 (bases 1 to 178184)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Oct 3, 2001 this sequence version replaced gi:9954580.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www-shgc.stanford.edu
 Quality: Phrap Quality >=40 99.7% of Sequence;
 Estimated Total Number of Errors is 2.5.

STS Content:

SHGC-37302 G30664
 SHGC-36253 G28435
 SHGC-14333 G13850.

FEATURES
 source
 1..178184
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTD-2006C1"

ORIGIN

Alignment Scores:
 Pred. No.: 3.67e-25 Length: 178184
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 8 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x AC008770 (1-178184)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
 Db 158632 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTCGCTCACACCTTCAAAATGATGGAAGGACT 158573

Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 158572 CACACTGGAGAGAAACCCATGAATGTAAGCAGTGTGGAGTCTTTGGATGTGCTCG 158513

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyLysGlnCys 60
 Db 158512 CGACTTCAAAATGATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 158453

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 158452 GGGAAAGCTTTGGATGTCCCTCAAACTTCGAAGCATGGAAGGACTCAC 158402

RESULT 11
 AC009397/c 179581 bp DNA linear HTG 03-AUG-2000
 LOCUS Homo sapiens chromosome 7 clone RP11-197P13, WORKING DRAFT
 DEFINITION SEQUENCE, 3 unordered pieces.
 AC009397
 AC009397.7 GI:9665193
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 179581)
 AUTHORS Olson,M.V.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 179581)
 AUTHORS Bubb,K.L., Desmarais,C.L., Ramsey,S.A. and Hubley,R.M.
 TITLE Direct Submission
 JOURNAL Submitted (20-AUG-1999) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
 COMMENT On Aug 3, 2000 this sequence version replaced gi:8050913.
 Sequence Quality Assessment:
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

----- Genome Center
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: http://www.genome.washington.edu
 Contact: uwgchgs@u.washington.edu
 ----- Project Information
 Center project name: HsaChr7
 Center clone name: RP11-197P13 (UWGC:djs111)
 ----- Summary Statistics
 Sequencing vector: M13; X02513; 100% of reads
 Chemistry: Dye-primer-amersham; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 177991 bases at least Q40
 Consensus quality: 178749 bases at least Q30

Consensus quality: 179128 bases at least Q20
Insert size: 210180; 20.5% error; agarose-fp
Insert size: 179518; sum-of-contigs
Quality coverage: 7.36x in Q20 bases; agarose-fp
Quality coverage: 8.62x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 34830: contig of 34830 bp in length
* 34831 34930: gap of unknown length
* 34931 94552: contig of 59622 bp in length
* 94552 94652: gap of unknown length
* 94653 179581: contig of 84929 bp in length.

FEATURES

source

1. .179581
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/clone="RP11-197P13"
/sex="male"
/tissue_type="blood"
/clone_lib="RP11 human BAC library 11"

misc_feature

1. .34830
/notes="assembly_name:Contig61
clone_end:T7
vector_side:right"

gap

34831..34930
/estimated_length=unknown

misc_feature

34931..94552
/notes="assembly_name:Contig62"

gap

94553..94652
/estimated_length=unknown

misc_feature

94653..179581
/notes="assembly_name:Contig63
clone_end:SP6
vector_side:left"

ORIGIN

Alignment Scores:
Pred. No.: 3,71e-25 Length: 179581
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC009397 (1-179581)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
|||||
Db 115608 TGTAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 115549
|||||

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
|||||
Db 115548 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAGTCTTTGGATGTCCTCG 115489
|||||

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
|||||
Db 115488 CGACTTCAATGATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 115429
|||||

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
|||||
Db 115428 GGGAAAGCTTTTGGATGTCCTTCAACCTTCGAAGCATGGAAGGACTCAC 115378
|||||

RESULT 12

AC022415/c

LOCUS AC022415 180510 bp DNA linear PRI 22-DEC-2000
DEFINITION Homo sapiens chromosome 19 clone CTC-359D24, complete sequence.
AC022415
VERSION AC022415.5 GI:11968305
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 180510)
DOE Joint Genome Institute and Stanford Human Genome Center.

AUTHORS Direct Submission

TITLE Unpublished

REFERENCE 2 (bases 1 to 180510)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (03-FEB-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 180510)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell

Drive, Walnut Creek, CA 94598, USA

COMMENT On Dec 22, 2000 this sequence version replaced gi:7704995.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www.shgc.stanford.edu

Quality: Phrap Quality >=40 99.9% of Sequence;

Estimated Total Number of Errors is 0.1.

STS Content:

SHGC-6150 GI4148.

Location/Qualifiers

1..180510

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="19"

/clone="CTC-359D24"

ORIGIN

Alignment Scores:
Pred. No.: 3,73e-25 Length: 180510
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 8 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC022415 (1-180510)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
|||||
Db 8022 TGTAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 7963
|||||

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
|||||
Db 7962 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAGTCTTTGGATGTCCTCG 7903
|||||

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
|||||
Db 7902 CGACTTCAATGATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 7843
|||||

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
|||||
Db 7842 GGGAAAGCTTTTGGATGTCCTTCAACCTTCGAAGCATGGAAGGACTCAC 7792
|||||

RESULT 13

AX202430

LOCUS

DEFINITION

AX202430

AX202430

AX202430

AX202430

AX202430

AX202430

AX202430

```

VERSION      AX202430.1  GI:15392178
KEYWORDS     .
SOURCE       synthetic construct
ORGANISM     synthetic construct
REFERENCE    other sequences; artificial sequences.
1
AUTHORS      Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE        Methods and compositions to modulate expression in plants
JOURNAL      Patent: WO 0152620-A 18 26-JUL-2001;
              The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
              DISCOVERY, INC. (CA)
FEATURES     Location/Qualifiers
              1..3300
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:34630"
               /note="Partial sequence of pMal-Ap3 and zinc finger protein
               ZFPap3"
ORIGIN
Alignment Scores:
Pred. No.:      7,178-27      Length:      3300
Score:          325.00      Matches:      54
Percent Similarity: 84.4%      Conservative: 11
Best Local Similarity: 70.13%      Mismatches: 12
Query Match:     73.53%      Indels:      0
DB:              6          Gaps:      0

US-10-669-861-2_COPY_33_109 (1-77) x AX202430 (1-3300)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||  :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 2758 TGTCGGGAATGTTGTAAGTCTTCAGCCAGCAGCTCCTGGTGGCGCCAGCAGGTACC 2817

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||  :::::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 2818 CACACGGGTGAAACCCGTATAAATGCCCAGAGTGGCGCAAAATCTTTAGCCAGTCCAGC 2877

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
    |||  :::::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 2878 AACCTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGT 2937

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||  :::::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 2938 GGCAGTCTTTCAGCCAGTCCAGCACTGGTGGCGCCACCAACGACTACTCAC 2988

RESULT 14
AC103247/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-148G24, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC103247
AC103247.5 GI:30578854
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 216979)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divyak,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,T., Guarva,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., McNeill,T.Z., Meenen,E.,
 Mathiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Muidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shateman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,B., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.
 Direct Submission

Unpublished
 2 (bases 1 to 216979)
 Worley,K.C.
 Direct Submission
 Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 216979)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:23120050.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center of Medicine
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GJPH
 Center clone name: CH230-148G24
 ----- Summary Statistics
 Assembly program: Atlas 3.0;

Consensus quality: 211343 bases at least Q40
 Consensus quality: 212467 bases at least Q30
 Consensus quality: 213063 bases at least Q20
 Estimated insert size: 219835; sum-of-contigs estimation
 Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

*** NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 55572: contig of 55572 bp in length
 * 55573 55672: gap of unknown length
 * 55673 215556: contig of 159884 bp in length
 * 215557 215656: gap of unknown length
 * 215657 216979: contig of 1323 bp in length.

FEATURES
 source
 1. .216979

/organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-148G24"

misc_feature
 50131..50826
 /notes="clone_boundary"
 clone_end:17
 site:EcoRI

gap
 end_sequence:RWBAN48TJB"
 55573..55672
 /estimated_length=unknown
 215557..215656
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 1-7e-24 Length: 216979
 Score: 322.00 Matches: 55
 Percent Similarity: 85.71% Conservative: 11
 Best Local Similarity: 71.43% Mismatches: 11
 Query Match: 72.85% Indels: 0
 DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC103247 (1-216979)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 |||||
 Db 2995 TCGGTRGACTGTGCCAAAGCCTTCGGCGACCGCTCAGCCCTTATTCGGCACCAGAGAACA 2936
 |||||
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 |||||
 Db 2935 CACACTGGAGAGAGCCTTATAAATGCAAGGACTGTGGAAGAGCTTTTCAGCCAGAGCTCA 2876
 |||||
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 |||||
 Db 2875 TCTCTTACAAACACACAGAAACTCAGCTGAGAAAGACCTTAAAGTGAAGAATGT 2816
 |||||
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 |||||
 Db 2815 CGAAAAGCGCTTTAGCCAGAGTTTCATCTCTTCTCAACATCAGAAAATCAT 2765
 |||||

RESULT 15
 AC133610/c
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-70117, *** SEQUENCING IN PROGRESS
 ,, 3 unordered pieces.
 ACCESSION
 AC133610
 VERSION
 AC133610.2 GI:25073277
 KEYWORDS
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS

1 (bases 1 to 238307)
 Muzny, D., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhsuwa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokelemeh, O., Okwuonu, G., Olampunsgoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, X., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajis, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished

TITLE
 JOURNAL

2 (bases 1 to 238307)
 Rat Genome Sequencing Consortium.
 Direct Submission

REFERENCE
 AUTHORS

Submitted (16-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS

3 (bases 1 to 238307)
 Rat Genome Sequencing Consortium.
 Direct Submission

TITLE
 JOURNAL

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 19, 2002 this sequence version replaced gi:22901951.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KCIN

Center clone name: CH230-70117

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 201502 bases at least Q40

Consensus quality: 204933 bases at least Q30

Consensus quality: 207525 bases at least Q20

Estimated insert size: 207494; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 109488: contig of 109488 bp in length
* 109489 109588: gap of unknown length
* 109589 237009: contig of 127421 bp in length
* 237010 237109: gap of unknown length
* 237110 238307: contig of 1198 bp in length.

FEATURES

source
1..238307
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-70117"
40706..41956
/note="wgs_contig"
106919..109488
/note="wgs_contig"
109489..109588
/estimated_length=unknown
109589..110614
/note="wgs_contig"
237010..237109
/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 1.89e-24 Length: 238307
Score: 322.00 Matches: 55
Percent Similarity: 85.71% Conservative: 11
Best Local Similarity: 71.43% Mismatches: 11
Query Match: 72.85% Indels: 0
DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC133610 (1-238307)

Qy 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 2995 TCGGTGACTGTGCCAAGCCCTTCGGCACCGCTCAGCCCTATTTCGGCACGAGAACA 2936
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40

Db 2935 CACACTGGAGAGAAGCCCTTATAAATGCAAGGACTGTGGAAGGCTTTTCAGCCAGAGCTCA 2876
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
Db 2875 TCTCTTACAAACACACAGAAAACCTCACACTGGAGAAAGACCCCTATAAGTGTAAAGAATGT 2816
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2815 GGAAGAGCGTTTAGCCAGAGTTCACTCTCTCTCAACATCAGAAAACTCAT 2765

Search completed: January 12, 2006, 02:31:37
Job time : 3742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 00:35:34 ; Search time 462 Seconds
(without alignments)
1110.782 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CXCCKAFIQKSNLRHQRTH.....KCGKAFGCPNLRHRGRTH 77

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2_1/USPTO_spool/US10669861/runat_09012006_144452_20361/app_query.fasta_1.263
-DB=N_Geneseq -QFWT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10669861 @CN 1 1 727 @runat_09012006_144452_20361 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 21.*

1:	Geneseq1980s.*
2:	Geneseq1990s.*
3:	Geneseq2000s.*
4:	Geneseq2001as.*
5:	Geneseq2001bs.*
6:	Geneseq2002as.*
7:	Geneseq2002bs.*
8:	Geneseq2003as.*
9:	Geneseq2003bs.*
10:	Geneseq2003cs.*
11:	Geneseq2003ds.*
12:	Geneseq2004as.*
13:	Geneseq2004bs.*
14:	Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	1137	11 ADL33661	Adl33661 Zinc fing
2	329	74.4	3569	4 ABA07313	Abao7313 Human pan
3	329	74.4	3569	4 AAK90472	Aak90472 Human dig
4	327	74.0	672	6 ABK71555	Abk71555 Human dit

5	327	74.0	1668	5	AAS86929	Aas86929 DNA encod
6	327	74.0	1746	12	ADN99083	Adn99083 Novel hum
7	327	74.0	1746	12	ADO00652	Ado00652 Novel hum
8	327	74.0	1765	6	AAS18786	Aas18786 cDNA enco
9	327	74.0	2192	6	ABS51792	Abs51792 Human mdd
10	327	74.0	2217	8	ABX34635	Abx34635 Human mdd
11	327	74.0	2275	6	ABX34635	Abx34635 Human mdd
12	327	74.0	2280	11	ADM03740	Adm03740 Human nuc
13	327	74.0	2280	11	ADM03740	Adm03740 Human cDN
14	327	74.0	2718	12	ADN98793	Adn98793 Novel hum
15	327	74.0	2718	12	ADN98793	Adn98793 Novel hum
16	327	74.0	2718	12	ADN98793	Adn98793 Novel hum
17	327	74.0	3825	13	ADO00651	Ado00651 Novel hum
18	326	73.8	636	11	ADL33677	Adl33677 Zinc fing
19	325	73.5	3300	4	AAD11592	Aad11592 Partial s
20	321	72.6	330	4	AAD11613	Aad11613 Zinc fing
21	321	72.6	3300	4	AAD11588	Aad11588 Partial s
22	321	72.6	3300	4	AAD11589	Aad11589 Partial s
23	319	72.2	2056	4	AAK53064	Aak53064 Human pol
24	319	72.2	2175	5	AAS82573	Aas82573 DNA encod
25	319	72.2	2496	4	AAK52080	Aak52080 Human pol
26	319	72.2	2667	11	ADM02541	Adm02541 Human cDN
27	318	71.9	505	12	ACH68852	Ach68852 Human gen
28	316	71.5	2115	9	ACA98974	Aca98974 cDNA enco
29	315	71.3	468	9	ACH15895	Ach15895 Human adu
30	315	71.3	630	11	ADL33669	Adl33669 Zinc fing
31	315	71.3	827	4	AAH05264	Aah05264 Human cDN
32	315	71.3	1258	12	ACH92086	Ach92086 Human gen
33	315	71.3	1449	4	AAH15787	Aah15787 Human cDN
34	315	71.3	1512	4	AAI185395	Aai185395 Human pol
35	315	71.3	1890	10	ADC30336	Adc30336 Human nov
36	315	71.3	1914	10	ADB63310	Adb63310 Human cDN
37	315	71.3	2257	9	ACA98925	Aca98925 cDNA enco
38	315	71.3	2664	11	ADM01826	Adm01826 Human cDN
39	315	71.3	2890	12	ADQ25219	Adq25219 Human sof
40	315	71.3	3309	8	ACC46347	Acc46347 Human dit
41	315	71.3	3639	5	AAS64586	Aas64586 DNA encod
42	315	71.3	3832	13	ADR07945	Adr07945 Full leng
43	315	71.3	3839	6	ABK83826	Abk83826 Human cDN
44	315	71.3	3839	12	ADN04207	Adn04207 Antipsori
45	315	71.3	3839	13	ADR25049	Adr25049 Breast ca

ALIGNMENTS

RESULT 1
ADL33661
ID ADL33661 standard; DNA; 1137 BP.
XX
AC ADL33661;
XX
DT 20-MAY-2004 (first entry)
XX
DE Zinc finger domain-containing protein (Neurol-p65) coding sequence.
XX
KW modified cell; artificial transcription factor; stress resistance;
XX phenotypic trait alteration; zinc finger domain; gene; ds.
XX Unidentified.
XX
XX WO2003048345-A1.
XX
PD 12-JUN-2003.
XX
PF 07-DEC-2002; 2002WO-KR002309.
XX
PR 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
XX (TOOL-) TOOLGEN INC.
XX

PI Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
XX WPI; 2003-513760/48.
DR P-PSDB; ADL33662.
XX
XX New modified cell comprising a heterologous nucleic acid encoding an
PT artificial transcription factor that confers stress resistance, useful
PT for altering a phenotypic trait of a cell or organism.
XX
XX Disclosure; SEQ ID NO 201; 169pp; English.
XX
CC The invention comprises a modified cell containing a heterologous nucleic
CC acid encoding an artificial transcription factor that confers stress
CC resistance to the modified cell. The modified cell of the invention is
CC useful for altering a phenotypic trait of a cell or organism. The present
CC DNA sequence encodes a protein which contains zinc finger domains.
XX
SQ Sequence 1137 BP; 280 A; 373 C; 269 G; 215 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.99e-45 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps:

US-10-669-861-2_COPY_33_109 (1-77) x ADL33661 (1-1137)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 97 TGTAAAGATTGGGGAAGACTTTTCATTGAGAGTCAAAACCTCATCAGACACAGAACT 156

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 157 CACACCGGGGAAAAACCGTACAAAGTGTGAAGATGTGGCAAGCTTTTACCCAACTCTCA 216

QY 41 AsnLeuThrLysHisLysLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 217 AACCTTACTAACAATAAGAAAATTCATACCGGGGAAAAACCGTATAAATGTAAGCAATGT 276

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 277 GGGAAAGCTTTTGGATGTCCTCAAACTTCGAAGGATGAAGACTCAC 327

RESULT 2
ID ABA07313
XX ABA07313 standard; DNA; 3569 BP.
AC ABA07313;
XX
DT 14-JAN-2002 (first entry)
XX
DE Human pancreatic cancer related genomic DNA, SEQ ID NO: 632.
XX
KW Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
KW antihormone; antiulcer; thyroid-active; gene therapy; antisense therapy;
KW pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
KW diabetes; endocrine disorder; acromegaly; hyperthyroidism;
KW gastrointestinal disorder; Crohn's disease; duodenal ulcer; ds.
XX
OS Homo sapiens.
XX
PN WO200155206-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0001353.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR

PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

CC The invention relates to an isolated polypeptide comprising an amino acid
 CC sequence at least 90% identical to 188 amino acid sequences fully defined
 CC in the specification and encoded by 188 cDNA clones fully defined in the
 CC specification. The invention also relates to a fragment having biological
 CC activity, a domain, an epitope, full length protein, variant, allelic
 CC variant or a species homologue of the fully defined sequence. The
 CC polynucleotide and polypeptide are useful for treating, preventing and/or
 CC prognosing disorders related to the pancreas including pancreatic cancer,
 CC pancreatitis, diabetes, endocrine disorders such as acromegaly or
 CC hyperthyroidism, and gastrointestinal disorders such as Crohn's disease
 CC and duodenal ulcers. The present sequence encodes a pancreatic cancer-
 CC related polypeptide of the invention
 XX
 SQ Sequence 3569 BP; 1193 A; 673 C; 773 G; 929 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1.32e-30 Length: 3569
 Score: 329.00 Matches: 54
 Percent Similarity: 84.42% Conservative: 11
 Best Local Similarity: 70.13% Mismatches: 12
 Query Match: 74.43% Indels: 0
 DB: 4 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x ABA07313 (1-3569)
 QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 DB 2436 TGCAGCAATGTGGTAAGCCCTTCATTTCATTCAGTCCCTCGATATCATGAAGGACT 2495
 QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 DB 2496 CACACTGGAGAGAAACCCCTATGAGTGAAGCAATGTGGAGAGCCCTTCAGATCTGCCTCA 2555
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 DB 2556 CACCTTCGAAATCGTAAAGGACTCACACTGGAGAGAAACCCCTATGAGTGAAGCAATGT 2615
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 DB 2616 GGGAAAGCCTTCAGTGTGCTCAACCTTCGAAAGCATGGTAGACTCAC 2666
 RESULT 3
 AAK90472
 ID AAK90472 standard; DNA; 3569 BP.
 XX AAK90472;
 AC AAK90472;
 XX
 DT 05-NOV-2001 (first entry)
 XX Human digestive system antigen genomic sequence SEQ ID NO: 4048.
 DE
 XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum; ds.
 XX Homo sapiens.
 XX
 XX WO20015314-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001324.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.

PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254907P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-457717/49.
 XX
 XX Isolated pancreatic cancer polypeptide for treating, preventing and/ or
 XX prognosing disorders related to the pancreas including pancreatic cancers
 XX and also for testing and detection e.g. diagnosis.
 XX
 XX Disclosure; SEQ ID NO 632; 537pp; English.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0233080P.
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233297P.
PR 14-SEP-2000; 2000US-0233298P.
PR 14-SEP-2000; 2000US-0233299P.
PR 14-SEP-2000; 2000US-023400P.
PR 14-SEP-2000; 2000US-023401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 4048; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention

```
XX SQ Sequence 3569 BP; 1193 A; 673 C; 773 G; 929 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 1-32e-30 Length: 3569
Score: 329.00 Matches: 54
Percent Similarity: 84.42% Conservative: 11
Best Local Similarity: 70.13% Mismatches: 12
Query Match: 74.43% Indels: 0
DB: 4 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x AAK90472 (1-3569)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 2436 TGCAGCAATGTGGTAAAGCCTTCATTCCAGTTCCTTCGATATCATGAAGGACT 2495
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 2496 CACACTGGAGAGAAACCCCTATGAGTGTAAAGCAATGTGGGAAGCCCTTCAGATCTGCCTCA 2555
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2556 CACCTTCGAATGCATGAAGGACTCACACTGGAGAGAAACCCCTATGAGTGAAGCAATGT 2615
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2616 GGGAAAGCCTTCAGTGTGCCTCAACCTTCGAAGCAATGTGAGGACTCAC 2666
RESULT 4
ID ABK71555 standard; cDNA; 672 BP.
XX AC ABK71555;
XX DT 30-JUL-2002 (first entry)
XX DE Human dithp polynucleotide #21.
XX KW Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone;
XX KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
XX KW inflammatory disorder; viral infection; bacterial infection; seizure;
XX KW fungal infection; parasitic infections; developmental disorder; breast;
XX KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
XX KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
XX KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
XX KW skin; testis; thymus.
OS Homo sapiens.
XX WO200220754-A2.
XX PD 14-MAR-2002.
XX PF 29-AUG-2001; 2001WO-US027127.
XX PR 05-SEP-2000; 2000US-0229747P.
XX PR 05-SEP-2000; 2000US-0229748P.
XX PR 05-SEP-2000; 2000US-0229749P.
XX PR 05-SEP-2000; 2000US-0229750P.
XX PR 05-SEP-2000; 2000US-0229751P.
XX PR 05-SEP-2000; 2000US-0230583P.
XX PR 06-SEP-2000; 2000US-0230505P.
XX PR 06-SEP-2000; 2000US-0230514P.
XX PR 06-SEP-2000; 2000US-0230515P.
XX PR 06-SEP-2000; 2000US-0230517P.
XX PR 06-SEP-2000; 2000US-0230518P.
XX PR 06-SEP-2000; 2000US-0230519P.
XX PR 06-SEP-2000; 2000US-0230595P.
XX PR 06-SEP-2000; 2000US-0230597P.
XX PR 06-SEP-2000; 2000US-0230598P.
XX PR 06-SEP-2000; 2000US-0230599P.
XX PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230865P.
PR 06-SEP-2000; 2000US-0230988P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-0231167P.
XX (INCY-) INCVTE GENOMICS INC.
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JU;
XX Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
XX Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
XX Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
XX Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI: 2002-383054/41.
XX P-PSDB; ABG59963.
XX An isolated polynucleotide useful in diagnostics and therapeutics.
XX Claim 1; Page 415; 686pp; English.
XX The invention relates to human diagnostic and therapeutic (dithp)
XX polynucleotides and their associated polypeptides (DITHP polypeptides).
XX The sequences of the invention are used in the treatment and diagnosis of
XX cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
XX (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
XX cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
XX thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
XX psoriasis, osteoporosis), viral infections, bacterial infections, fungal
XX infections, parasitic infections, developmental disorders (e.g. anaemia,
XX epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
XX endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
XX (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
XX amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
XX disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
XX (e.g. myotonic dystrophy, catatonia, peripheral neuropathy). Sequences
XX ABK71535-ABK71809 represent human dithp polynucleotides of the invention
XX SQ Sequence 672 BP; 213 A; 137 C; 163 G; 159 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.56e-31 Length: 672
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.72% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x ABK71555 (1-672)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 311 TGTAAAGCAGTGTGGAAAGCCCTTCGATCTGCCTCACACCTTCAATGCATGAAGGACT 370
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 371 CACACTGGAGAGAAACCCCTATGATGTAAAGCAGTGTGGAGAGCTTTGGATGTCCTCG 430
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 431 CGACTTCAATGCATGAAGGACTCACACTGGAGAGAAACCCCTATATAAATGAAGCAATCT 490
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 491 GGGAAAGCCTTTTGGATGTCCCTCAACCTTCGAAGGCATGAAGGACTCAC 541
RESULT 5
ID AAS86929 standard; cDNA; 1668 BP.
XX AC AAS86929;
XX DT 13-FEB-2002 (first entry)
```

XX DNA encoding novel human diagnostic protein #22733.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG22742.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 22733; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (III) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1668 BP; 555 A; 334 C; 389 G; 390 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,588-31 Length: 1668
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 5 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AAS86929 (1-1668)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 595 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 654

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrClnSerSer 40
DB 655 CACACTGGAGAGAAACCCATATGATGAAGCAGTGTGGGAAAGTCTTTGGATGTCCCTCG 714

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 715 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCGTATATAATGTAAGCAATGT 774

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
DB 775 GGGAAAGCTTTGGATGTCCCTCAAAAGCTTCGAAGGCATGGAAGGACTCAC 825

RESULT 6
ADN99083
ID ADN99083 standard; cDNA; 1746 BP.
XX AC ADN99083;
XX DT 29-JUL-2004 (first entry)
XX DE Novel human cDNA sequence #683.
XX ds; gene; anti-inflammatory; dermatological; neuroprotective;
KW immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;
KW gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
KW psoriasis; diabetes; early aging; hormonal imbalance;
KW ischemic heart disease; ulcerative colitis.
XX Homo sapiens.
PN WO2004038003-A2.
XX 06-MAY-2004.
XX 24-OCT-2003; 2003WO-US033947.
XX 25-OCT-2002; 2002US-0421061P.
PR 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 30-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426355P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429302P.
PR 27-NOV-2002; 2002US-0429326P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433501P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-04711306P.
 19-MAY-2003; 2003US-04711336P.
 22-MAY-2003; 2003US-0472420P.
 22-MAY-2003; 2003US-0472430P.
 09-JUN-2003; 2003US-0476609P.
 09-JUN-2003; 2003US-0476621P.
 09-JUN-2003; 2003US-0476632P.
 09-JUN-2003; 2003US-0476641P.
 08-JUL-2003; 2003US-0485217P.
 08-JUL-2003; 2003US-0485218P.
 08-JUL-2003; 2003US-0485223P.
 08-JUL-2003; 2003US-0485224P.
 08-JUL-2003; 2003US-0485325P.
 08-JUL-2003; 2003US-0485359P.
 14-JUL-2003; 2003US-0486446P.
 14-JUL-2003; 2003US-0486480P.
 15-JUL-2003; 2003US-0486891P.
 15-JUL-2003; 2003US-0486960P.
 08-AUG-2003; 2003US-0493341P.
 08-AUG-2003; 2003US-0493370P.
 08-AUG-2003; 2003US-0493573P.
 08-AUG-2003; 2003US-0493577P.
 (FIVE-) FIVE PRIME THERAPEUTICS INC.
 Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
 Wong JGP, Wu G, Zhang H, Zeng C;
 WPI; 2004-365511/34.
 P-PSDB; ADM99867.
 New nucleic acid molecules, useful in preparing a composition for
 treating or preventing e.g. inflammatory CNS, bacterial or viral
 disorders, cancer, psoriasis, diabetes, early aging, ischemic heart disease or
 ulcerative colitis.
 Claim 1; SEQ ID NO 683; 532pp; English.
 The invention relates to a nucleic acid molecule comprising a
 polynucleotide sequence or its complement that encodes a polypeptide. The
 nucleic acid is useful in preparing a composition for treating or
 preventing inflammatory, CNS, immune, bacterial or viral disorder,
 cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
 heart disease or ulcerative colitis. This sequence corresponds to a
 nucleic acid of the invention.
 Sequence 1746 BP; 595 A; 349 C; 374 G; 428 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9,12e-31 Length: 1746
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 12 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x ADM99083 (1-1746)
 QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 1414 TGTAAAGCAGTGTGGGAAGCCCTTGGATCGCTCACACCTTCAAAATGCATGGAAGGACT 1473
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 1474 CACATGGAGAGAACCCCTATGAAGTAAAGCAGTGGGAAGTCTTGGATGTGCCTCG 1533
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 1534 CGACTTCAATGCATGGAAGGACTCACATGGAGAGAACCCCTATGAAGTAAATGTAAGCAATGT 1593
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 DB: 1593

Db 1594 GGAAAGCTTTGGATGTCCTCAAACTTCGAAGGCTGGAAGGACTCAC 1644
 RESULT 7
 ADO00652
 ID ADO00652 standard; cDNA; 1746 BP.
 XX
 AC ADO00652;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Novel human cDNA sequence #1467.
 XX
 KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
 immunomodulator; antibacterial; virucide; antiproliferative; cytostatic;
 gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
 psoriasis; diabetes; early aging; hormonal imbalance;
 ischemic heart disease; ulcerative colitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2004038003-A2.
 XX
 PD 06-MAY-2004.
 XX
 PF 24-OCT-2003; 2003WO-US033947.
 XX
 PR 25-OCT-2002; 2002US-0421061P.
 PR 25-OCT-2002; 2002US-0421080P.
 PR 25-OCT-2002; 2002US-0421552P.
 PR 25-OCT-2002; 2002US-0421614P.
 PR 30-OCT-2002; 2002US-0422177P.
 PR 30-OCT-2002; 2002US-0422178P.
 PR 15-NOV-2002; 2002US-0426355P.
 PR 15-NOV-2002; 2002US-0426384P.
 PR 15-NOV-2002; 2002US-0426394P.
 PR 15-NOV-2002; 2002US-0426430P.
 PR 15-NOV-2002; 2002US-0426916P.
 PR 27-NOV-2002; 2002US-0429224P.
 PR 27-NOV-2002; 2002US-0429275P.
 PR 27-NOV-2002; 2002US-0429302P.
 PR 27-NOV-2002; 2002US-0429326P.
 PR 04-DEC-2002; 2002US-0429651P.
 PR 04-DEC-2002; 2002US-0430645P.
 PR 04-DEC-2002; 2002US-0430651P.
 PR 04-DEC-2002; 2002US-0430657P.
 PR 04-DEC-2002; 2002US-0430663P.
 PR 04-DEC-2002; 2002US-0430668P.
 PR 04-DEC-2002; 2002US-0430684P.
 PR 05-DEC-2002; 2002US-0430937P.
 PR 05-DEC-2002; 2002US-0430965P.
 PR 05-DEC-2002; 2002US-0431458P.
 PR 12-DEC-2002; 2002US-0433251P.
 PR 12-DEC-2002; 2002US-0433500P.
 PR 13-DEC-2002; 2002US-0433316P.
 PR 13-DEC-2002; 2002US-0433318P.
 PR 23-DEC-2002; 2002US-0436236P.
 PR 03-JAN-2003; 2003US-0437914P.
 PR 17-JAN-2003; 2003US-0440820P.
 PR 17-JAN-2003; 2003US-0440821P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471338P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476621P.

```
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RP, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
XX P-PSDB; ADN99867.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 1; SEQ ID NO 2251; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX
XX Sequence 1746 BP; 595 A; 349 C; 374 G; 428 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. NO.: 9,128-31 Length: 1746
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x ADO0652 (1-1746)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
DB 1414 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1473
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1474 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAAGTCTTTGGATGTCCTCG 1533
QY 41 AsnLeuThrLysHisLysAlaPheIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
DB 1534 CGACTTCAATGTCGGAAGGACTCACACTGGAGAGAAACCCGTATGAATGTAAGCAATGT 1593
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
DB 1594 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAAGGCATGGAGACTCAC 1644
RESULT 8
ID AAS18786 standard; cDNA; 1765 BP.
XX
```

```
AC AAS18786;
XX
XX 26-MAR-2002 (first entry)
XX
XX cDNA encoding human zinc finger protein 53.
XX
XX Human; zinc finger protein 53; malignant tumour; HIV infection;
XX immunological disease; inflammation; 88.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 189..1640
XX FT /*tag= a
XX FT /product= "Zinc finger protein 53"
XX
XX CNI307009-A.
XX
XX 08-AUG-2001.
XX
XX 21-JAN-2000; 2000CN-00111472.
XX
XX 21-JAN-2000; 2000CN-00111472.
XX (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX WPI; 2002-049871/07.
XX P-PSDB; AAU10794.
XX
XX Polypeptide-human zinc finger protein 53 and polynucleotide for coding
XX said polypeptide.
XX
XX Claim 6; Page 25-26 (Disclosure); 37pp; Chinese.
XX
XX The present invention relates to the isolation of human zinc finger
XX protein 53 and the cDNA sequence encoding it. Also described is the DNA
XX recombination process to produce zinc finger protein 53, and methods of
XX applying zinc finger protein 53 and the polynucleotide encoding it in
XX treating various diseases, such as malignant tumours, HIV infection,
XX immunological diseases and inflammations. The present invention also
XX discloses the antagonist resisting the polypeptide and its treatment
XX effect. The present sequence encodes for the human zinc finger protein 53
XX of the invention
XX
XX Sequence 1765 BP; 597 A; 354 C; 383 G; 431 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. NO.: 9,258-31 Length: 1765
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x AAS18786 (1-1765)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
DB 1308 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1367
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1368 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAAGTCTTTGGATGTCCTCG 1427
QY 41 AsnLeuThrLysHisLysAlaPheIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
DB 1428 CGACTTCAATGTCGGAAGGACTCACACTGGAGAGAAACCCGTATGAATGTAAGCAATGT 1487
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
DB 1488 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAAGGCATGGAGACTCAC 1538
```


RESULT 9
ABSS1792
ID ABSS1792 standard; cDNA; 2192 BP.
XX
AC ABSS1792;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human mddt cDNA Incyte ID No: LI:1175131.1:2001JAN12.
XX
KW Human; molecule for disease detection and treatment; MDDT; cancer;
KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiasthmatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200255738-A2.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2002; 2002WO-US001008.
XX
PR 12-JAN-2001; 2001US-0261622P.
PR 16-JAN-2001; 2001US-0261865P.
PR 17-JAN-2001; 2001US-0262208P.
PR 17-JAN-2001; 2001US-0262209P.
PR 17-JAN-2001; 2001US-0262326P.
PR 19-JAN-2001; 2001US-0263063P.
PR 19-JAN-2001; 2001US-0263065P.
PR 19-JAN-2001; 2001US-0263329P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
PI Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;
XX
DR WPI; 2002-590679/63.
DR P-PSDB; ABG70319.
XX
XX New disease detection and treatment molecule (MDDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDDT expression, e.g. autoimmune or inflammatory
PT disorders.
XX
XX Claim 1; Page 93; 129pp; English.
XX
XX The present invention relates to the isolation of novel human molecules
CC for disease detection and treatment (MDDT), and the polynucleotide
CC sequences (mddt) encoding them. The MDDT polypeptides may be used to
CC screen for molecules that bind to, or are bound by the encoded
CC polypeptides, and to develop a transcript image of a tissue or cell type.
CC Probes comprising at least 20 nucleotides of the mddt polynucleotide may
CC be used to assess the toxicity of a test compound. The MDDT polypeptides
CC and mddt polynucleotides are useful in the diagnosis, study, prevention
CC and treatment of diseases associated with the expression of molecules for
CC disease detection and treatment. Such disorders include cell
CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, or cancers),
CC and autoimmune/inflammatory disorders (e.g. asthma, Crohn's disease, or
CC multiple sclerosis). The mddt polynucleotides may also be used as
CC molecule markers, in microarrays, and in somatic or germline gene
CC therapy. ABSS1779-ABSS1814 encode the MDDT proteins of the invention
XX
SQ Sequence 2192 BP; 709 A; 450 C; 503 G; 528 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 1,236-30 Length: 2192
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14

Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x ABSS1792 (1-2192)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 1845 TGTAAAGCAGTGTGGGAAAGCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1904
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGluGluCysGluGluGluGluGluGluGlu 40
DB 1905 CACACTGGAGAGAAACCCCTATCAATGTAAGCAGTGTGGGAAAGTCTTTTGGATGTCCTG 1964
QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 1965 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCCCTATCAATGTAAGCAGTGT 2024
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 2025 GGGAAAGCTTTTGGATGTCCCTCAAACTTCGAAGGCATGGAAGGACTCAC 2075
RESULT 10
ABX34635
ID ABX34635 standard; cDNA; 2217 BP.
XX
AC ABX34635;
XX
DT 13-FEB-2003 (first entry)
XX
DE Human mddt cDNA SEQ ID 196.
XX
KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200279449-A2.
XX
PD 10-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-US009944.
XX
XX 28-MAR-2001; 2001US-0279619P.
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300001P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshey SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen AD, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR WPI; 2003-058431/05.
DR P-PSDB; ABU11645.
XX
PT New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis.

XX PS Claim 1; SEQ ID NO 196; 339pp + Sequence Listing; English.

XX CC This invention describes a novel disease detection and treatment molecule

XX CC polypeptide (MDDR) which has anti-inflammatory, immunosuppressive,

XX CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,

XX CC antianemic, antipsoriatic and hepatotropic activity. The polynucleotides

XX CC and the polypeptides of the invention can be used for gene therapy,

XX CC protein replacement therapy and are useful for treating a variety of

XX CC diseases or conditions. These polypeptides or polynucleotides are

XX CC particularly useful for diagnosing, treating or preventing cell

XX CC proliferative disorders (e.g. cancers including adenocarcinoma,

XX CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's

XX CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's

XX CC syndrome, inflammation, osteoporosis, thrombocytopaenia, psoriasis or

XX CC hepatitis. ABX34440-ABX34835 encode the MDDR polypeptides represented in

XX CC ABU11450-ABU11845, described in the disclosure of the invention. NOTE:

XX CC The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2217 BP; 711 A; 455 C; 521 G; 530 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.25e-30 Length: 2217

Score: 327.00 Matches: 56

Percent Similarity: 81.82% Conservative: 7

Best Local Similarity: 72.73% Mismatches: 14

Query Match: 73.98% Indels: 0

DB: 8 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ABX34635 (1-2217)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20

Db 1856 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATCATGGAAGGACT 1915

Qy 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40

Db 1916 CACACTGGAGAGAAACCTTATGAATGTAAGCAGTGTGGAAAGCTTTTGGATGTGCCCTCG 1975

Qy 41 AsnLeuThrLysHisLysLysAlaIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60

Db 1976 CGACTTCAAAATGCATGGAAGGACTCACACTGGAGAGAAACCTTATGAATGTAAGCAGTGT 2035

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77

Db 2036 GGGAAAGCTTTTGGATGTCCCTCAAAACCTTGAAGGATGGAAGGACTCAC 2086

RESULT 11

AAD41202

ID AAD41202 standard; cDNA; 2275 BP.

AC AAD41202;

XX 30-OCT-2002 (first entry)

XX Human nucleic acid-associated protein (NAAP-12) cDNA.

XX Human; nucleic acid-associated protein; NAAP-12; neurological disorder;

XX arteriosclerosis; cancer; cell proliferative disorder; arteriosclerosis;

XX lymphoma; epilepsy; Alzheimer's disease; developmental; anticonvulsant;

XX autoimmune disorder; AIDS; allergy; anaemia; stroke; malaria; leishmania;

XX gene therapy; neurotropic; neuroprotective; cerebroprotective; virucide;

XX immunosuppressive; protozoacide; antimicrobial; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 234..2246

XX /*tag= a

XX /product= "Human NAAP-12"

XX FT

XX ID ADM03740 standard; cDNA; 2280 BP.

XX

PN WO200250279-A2.

XX PD 27-JUN-2002.

XX 19-DEC-2001; 2001WO-US050256.

XX 21-DEC-2000; 2000US-0257714P.

XX 05-JAN-2001; 2001US-0260081P.

XX 16-JAN-2001; 2001US-0262302P.

XX 23-JAN-2001; 2001US-0263823P.

XX 02-FEB-2001; 2001US-0266088P.

XX 29-OCT-2001; 2001US-0348442P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughin MR, Lu Y, Arvizu C, Ramkumar J, Yao MG, Policky JL;

XX Walla NK, Tribouley KW, Yue H, Batra S, Ding L, Lal PG;

XX Borowsky ML, Lu DAM, Gandhi AR, Griffin JA, Xu Y, Azimzai Y;

XX Gietzen KJ, Tang VT, Warren BA, Mason PM, Burford N, Hafalia AJA;

XX Lee EA, Yang J, Gorvad AE, Emerling BM, Marquis JP, Lee SY;

XX Swarnakar A, Reddy R;

XX WPI; 2002-519887/55.

XX P-PSDB; AAE25293.

XX Nucleic acid associated proteins and nucleic acids for diagnosing,

XX treating and preventing cell proliferative (e.g. cancers), neurological

XX (e.g. epilepsy or stroke) or autoimmune disorders (e.g. AIDS).

XX Claim 83; Page 188; 193pp; English.

XX The invention relates to nucleic acid-associated proteins (NAAP) and

XX nucleic acids. The nucleic acid and amino acid sequences are useful for

XX diagnosing, treating and preventing cell proliferative e.g.

XX arteriosclerosis, atherosclerosis, lymphoma or cancers), neurological

XX (e.g. epilepsy, Alzheimer's disease or stroke), developmental, and

XX autoimmune disorders (e.g. AIDS, allergies, or anaemia) or infections

XX (e.g. malaria, or leishmania), as well as in assessing the effects of

XX exogenous compound on the expression of nucleic acid and amino acid

XX sequences of nucleic acid-associated proteins. The invention is useful in

XX gene therapy. The present sequence is human NAAP-12 cDNA

XX SQ Sequence 2275 BP; 716 A; 473 C; 548 G; 538 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.3e-30 Length: 2275

Score: 327.00 Matches: 56

Percent Similarity: 81.82% Conservative: 7

Best Local Similarity: 72.73% Mismatches: 14

Query Match: 73.98% Indels: 0

DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AAD41202 (1-2275)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20

Db 1914 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATCATGGAAGGACT 1973

Qy 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40

Db 1974 CACACTGGAGAGAAACCTTATGAATGTAAGCAGTGTGGAAAGCTTTTGGATGTGCCCTCG 2033

Qy 41 AsnLeuThrLysHisLysLysAlaIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60

Db 2034 CGACTTCAAAATGCATGGAAGGACTCACACTGGAGAGAAACCTTATGAATGTAAGCAGTGT 2093

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77

Db 2094 GGGAAAGCTTTTGGATGTCCCTCAAAACCTTGAAGGATGGAAGGACTCAC 2144

RESULT 12

ADM03740

ID ADM03740 standard; cDNA; 2280 BP.

PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485329P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
DR WPI; 2004-365511/34.
DR P-P5DB; ADN99577.
XX

PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX

PS Claim 1; SEQ ID NO 1961; 532pp; English.

XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX

SQ Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.646-30 Length: 2718
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADO00362 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1687 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGCATGGAAGGACT 1746
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1747 CACACTGGAGAGAACCCCTATGATGTAGACAGTGTGGAGAGTCTTTGGATGTGCCTCG 1806
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1807 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAACCCGTATATATGTAAGCAATGT 1866
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1867 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAGGCATGGAAGGACTCAC 1917

RESULT 14

ADN98793

ID ADN98793 standard; cDNA; 2718 BP.

XX

AC ADN98793;

XX

DT 29-JUL-2004 (first entry)
XX Novel human cDNA sequence #393.
DE
XX
KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
KW immunomodulator; antibacterial; virucide; antipsoxiatic; cytostatic;
KW gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
KW psoriasis; diabetes; early aging; hormonal imbalance;
KW ischemic heart disease; ulcerative colitis.
XX
OS Homo sapiens.
XX
PN WO2004038003-A2.
XX
XX 06-MAY-2004.
PD
XX 24-OCT-2003; 2003WO-US033947.
PF 25-OCT-2002; 2002US-0421061P.
PF 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 25-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426355P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429302P.
PR 27-NOV-2002; 2002US-0429326P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430688P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 12-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-043251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 23-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.

PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
XX P-PSDB; ADN99577.
XX

XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX

PS Claim 1; SEQ ID NO 393; 532pp; English.

XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX

XX Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.64e-30 Length: 2718
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservatives: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 13.98% Indels: 0
DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADN98793 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
Db 1687 TGTAAAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1746
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1747 CACACTGGAGAGAACCCCTATGAATGTAAGCAGTGTGGGAAGCTTTTGGATGTCCTCG 1806
QY 41 AsnLeuThrLysHisLysLysLeuHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 1807 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAACCCCTATGAATGTAAGCAGTGT 1866
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1867 GGGAAAGCTTTTGGATGTCCTCAAACTTCGAAGCATGGAAGGACTCAC 1917

RESULT 15

ADN99082
ID ADN99082 standard; cDNA; 2718 BP.

XX AC

XX ADN99082;

XX DT 29-JUL-2004 (first entry)

XX DE Novel human cDNA sequence #682.

XX KW ds; gene; anti-inflammatory; dermatological; neuroprotective;

XX immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;

gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
psoriasis; diabetes; early aging; hormonal imbalance;
ischemic heart disease; ulcerative colitis.

Homo sapiens.

WO2004038003-A2.

06-MAY-2004.

24-OCT-2003; 2003WO-US033947.

25-OCT-2002; 2002US-0421061P.

25-OCT-2002; 2002US-0421080P.

25-OCT-2002; 2002US-0421552P.

25-OCT-2002; 2002US-0421614P.

30-OCT-2002; 2002US-0422177P.

30-OCT-2002; 2002US-0422178P.

15-NOV-2002; 2002US-0426355P.

15-NOV-2002; 2002US-0426384P.

15-NOV-2002; 2002US-0426394P.

15-NOV-2002; 2002US-0426430P.

15-NOV-2002; 2002US-0426916P.

27-NOV-2002; 2002US-0429224P.

27-NOV-2002; 2002US-0429275P.

27-NOV-2002; 2002US-0429302P.

27-NOV-2002; 2002US-0429328P.

04-DEC-2002; 2002US-0430651P.

04-DEC-2002; 2002US-0430645P.

04-DEC-2002; 2002US-0430651P.

04-DEC-2002; 2002US-0430657P.

04-DEC-2002; 2002US-0430663P.

04-DEC-2002; 2002US-0430668P.

05-DEC-2002; 2002US-0430684P.

05-DEC-2002; 2002US-0430937P.

05-DEC-2002; 2002US-0430965P.

12-DEC-2002; 2002US-0431458P.

12-DEC-2002; 2002US-043251P.

13-DEC-2002; 2002US-0433500P.

13-DEC-2002; 2002US-0433316P.

13-DEC-2002; 2002US-0433318P.

23-DEC-2002; 2002US-0436236P.

03-JAN-2003; 2003US-0437914P.

17-JAN-2003; 2003US-0440820P.

17-JAN-2003; 2003US-0440821P.

18-APR-2003; 2003US-0463700P.

18-APR-2003; 2003US-0463708P.

18-APR-2003; 2003US-0463716P.

18-APR-2003; 2003US-0463732P.

02-MAY-2003; 2003US-0467199P.

02-MAY-2003; 2003US-0467201P.

02-MAY-2003; 2003US-0467203P.

02-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-0471306P.

19-MAY-2003; 2003US-0471336P.

22-MAY-2003; 2003US-0472420P.

22-MAY-2003; 2003US-0472430P.

09-JUN-2003; 2003US-0476609P.

09-JUN-2003; 2003US-0476621P.

09-JUN-2003; 2003US-0476632P.

09-JUN-2003; 2003US-0476641P.

08-JUL-2003; 2003US-0485217P.

08-JUL-2003; 2003US-0485218P.

08-JUL-2003; 2003US-0485223P.

08-JUL-2003; 2003US-0485224P.

08-JUL-2003; 2003US-0485325P.

14-JUL-2003; 2003US-0485359P.

14-JUL-2003; 2003US-0486446P.

14-JUL-2003; 2003US-0486480P.

15-JUL-2003; 2003US-0486891P.

15-JUL-2003; 2003US-0486960P.

08-AUG-2003; 2003US-0493341P.

08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX WPI; 2004-365511/34.
DR P-PsDB; ADN99866.
XX
PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
PS Claim 1; SEQ ID NO 692; 532pp; English.
XX
CC The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX
SQ Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.64e-30 Length: 2718
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADN99082 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1687 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAATCGATGGAAGGACT 1746
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1747 CACACTGGAGAGAAACCTTATGAAATGTAGCAGTGTGGAAAGTCTTTGGATGTGCTCG 1806
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1807 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 1866
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1867 GGGAAAGCTTTTGGATGTCCCTCAAACCTTCGAGGCATGGAAGGACTCAC 1917

Search completed: January 12, 2006, 02:38:53
Job time : 467 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 01:29:19 ; Search time 155 Seconds

(without alignments)
883.047 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCGAFQKSLIRHQT.....KQCGAFGCPNLRHRGTH 77

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US1069861/runat_09012006.144453_20393/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1069861@cgn.1.1.193 @runat_09012006.144453_20393 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2.6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2.6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2.6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2.6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2.6/ptodata/1/ina/H COMB.seq:*
6: /cgn2.6/ptodata/1/ina/PCTUS COMB.seq:*
7: /cgn2.6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2.6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	315	71.3	601	3	US-09-949-016-27361 Sequence 27361, A
c 2	315	71.3	601	3	US-09-949-016-149636 Sequence 149636, A
c 3	315	71.3	1914	3	US-10-104-047-1464 Sequence 1464, Ap
4	315	71.3	3798	3	US-09-949-016-4204 Sequence 4204, Ap
5	315	71.3	3839	3	US-09-949-016-485 Sequence 485, App
6	315	71.3	156942	3	US-09-949-016-12227 Sequence 12227, A
7	315	71.3	156950	3	US-09-949-016-15946 Sequence 15946, A
8	314	71.0	3186	3	US-09-016-434-1390 Sequence 1390, Ap
9	313	70.8	1835	3	US-10-104-047-901 Sequence 901, App

10	309	69.9	601	3	US-09-949-016-27493 Sequence 27493, A
11	309	69.9	601	3	US-09-949-016-92521 Sequence 92521, A
12	309	69.9	1694	3	US-10-104-047-1500 Sequence 1500, Ap
13	309	69.9	1820	3	US-09-949-016-2645 Sequence 2645, Ap
14	309	69.9	1839	3	US-09-949-016-493 Sequence 493, App
15	309	69.9	2241	3	US-10-104-047-693 Sequence 693, App
16	309	69.9	23174	3	US-09-949-016-14387 Sequence 14387, A
17	309	69.9	23187	3	US-09-949-016-12335 Sequence 12335, A
18	308	69.7	2110	3	US-10-104-047-1778 Sequence 1778, Ap
19	307	69.5	2320	3	US-09-016-434-1054 Sequence 1054, Ap
20	307	69.5	4272	3	US-09-620-312D-586 Sequence 586, App
21	306	69.2	601	3	US-09-949-016-44799 Sequence 44799, A
22	306	69.2	2197	3	US-10-104-047-884 Sequence 884, App
23	305	69.0	2038	3	US-09-620-312D-911 Sequence 911, App
24	305	69.0	3078	3	US-10-104-047-622 Sequence 622, App
25	304	68.8	45138	3	US-09-949-016-13027 Sequence 13027, A
c 26	304	68.8	563	3	US-09-385-982-433 Sequence 433, App
27	302	68.3	1429	3	US-09-023-655-689 Sequence 689, App
28	302	68.3	1995	3	US-10-104-047-1730 Sequence 1730, Ap
29	301	68.1	2210	3	US-10-104-047-1453 Sequence 1453, Ap
30	301	68.1	3138	3	US-10-104-047-83 Sequence 83, Appl
31	299	67.6	2351	3	US-09-016-434-1337 Sequence 1337, Ap
32	299	67.6	2637	3	US-09-949-016-5823 Sequence 5823, Ap
33	299	67.6	3754	3	US-10-104-047-530 Sequence 530, App
34	299	67.6	4468	3	US-09-620-312D-243 Sequence 243, App
35	299	67.6	27227	3	US-09-949-016-17365 Sequence 17365, A
c 36	298	67.4	601	3	US-09-949-016-27359 Sequence 27359, A
c 37	298	67.4	601	3	US-09-949-016-27360 Sequence 27360, A
c 38	298	67.4	601	3	US-09-949-016-149637 Sequence 149637, A
c 39	298	67.4	601	3	US-09-949-016-149638 Sequence 149638, A
40	298	67.4	2441	3	US-09-949-016-2756 Sequence 2756, Ap
41	298	67.4	19861	3	US-09-949-016-14498 Sequence 14498, A
42	297	67.2	2487	3	US-10-104-047-898 Sequence 898, App
43	295	66.7	2779	3	US-09-976-594-175 Sequence 175, App
44	294	66.5	1460	3	US-09-774-528-413 Sequence 413, App
45	294	66.5	1460	3	US-10-120-988-413 Sequence 413, App

ALIGNMENTS

RESULT 1
US-09-949-016-27361/c
; Sequence 27361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27361
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27361

Alignment Scores:
Pred. No.: 2.74e-31 Length: 601
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27361 (1-601)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:::|||||
Db 274 TGTAAGAATGCGCAAGCATTTATATGGCTTCAACCCCTAACTAGACATAAGAGGATA 215
    |||||:::|||||
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:::|||||
Db 214 CACACTGGAGAGAGCCCTACAAATGTGAAGATGTGGCAAGCTTTTAGCGGTCTCTCA 155
    |||||:::|||||
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:::|||||
Db 154 ACCCTTACTAAGCATAAAGCAATTCATCTCGAGAGAAACCCCTACAAATGTAAAGAATGT 95
    |||||:::|||||
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:::|||||
Db 94 GGCAAAAGCTTTTAAGCACTCTCTCAGCCCTTGCTAAACATAAATAATATACAT 44
    |||||:::|||||
```

RESULT 2

US-09-949-016-149636/c
; Sequence 149636, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 149636

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-149636

Alignment Scores:

Pred. No.:	2,748-31	Length:	601
Score:	315.00	Matches:	56
Percent Similarity:	80.52%	Conservative:	6
Best Local Similarity:	72.73%	Mismatches:	15
Query Match:	71.27%	Indels:	0
DB:	3	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-149636 (1-601)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:::|||||
Db 274 TGTAAGAATGCGCAAGCATTTATATGGCTTCAACCCCTAACTAGACATAAGAGGATA 215
    |||||:::|||||
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:::|||||
Db 214 CACACTGGAGAGAGCCCTACAAATGTGAAGATGTGGCAAGCTTTTAGCGGTCTCTCA 155
    |||||:::|||||
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:::|||||
Db 154 ACCCTTACTAAGCATAAAGCAATTCATCTCGAGAGAAACCCCTACAAATGTAAAGAATGT 95
    |||||:::|||||
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:::|||||
Db 94 GGCAAAAGCTTTTAAGCACTCTCTCAGCCCTTGCTAAACATAAATAATATACAT 44
    |||||:::|||||
```

RESULT 3

US-10-104-047-1464

; Sequence 1464, Application US/10104047

; Patent No. 6943241

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. 6943241el full length cDNA

; FILE REFERENCE: HI-A0105

; CURRENT APPLICATION NUMBER: US/10/104,047

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER:

; PRIOR FILING DATE:

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1464

; LENGTH: 1914

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-104-047-1464

Alignment Scores:

Pred. No.:	1,35e-30	Length:	1914
Score:	315.00	Matches:	55
Percent Similarity:	77.92%	Conservative:	5
Best Local Similarity:	71.43%	Mismatches:	17
Query Match:	71.27%	Indels:	0
DB:	3	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-1464 (1-1914)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:::|||||
Db 1557 TGCAGTGAATGTGGAAAGCTTTTATTACAGATGTCAAACCTCATTCGACACAGAGAATT 1616
    |||||:::|||||
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:::|||||
Db 1617 CATACGGGTGAGAAACCCCTATGCATGTACATGTGTGAAAAGCCTTTAGTCAGAAAATCA 1676
    |||||:::|||||
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:::|||||
Db 1677 AACCTCACTGAACATGAGAAAATTCATCTCGAGAGAAACCTTATCATTTGTAATCAATGT 1736
    |||||:::|||||
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:::|||||
Db 1737 GGGAAAGCTTTTCAGTCAGAGACAAAATCTTCTTGCGATGAAAAAATTCAT 1787
    |||||:::|||||
```

RESULT 4

US-09-949-016-4204

; Sequence 4204, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4204

; LENGTH: 3798

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4204

Alignment Scores:

Pred. No.:	3,49e-30	Length:	3798
Score:	315.00	Matches:	56
Percent Similarity:	80.52%	Conservative:	6
Best Local Similarity:	72.73%	Mismatches:	15
Query Match:	71.27%	Indels:	0


```

DB: 3 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-4204 (1-3798)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 2368 TGTAAAGAAATGTGGCAAGCAATTATATGCTTCAACCTTACTACTAGATAGAGGATA 2427
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 2428 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 2487
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2488 ACCCTTACTAAGCATAAGACAATTCTACTGGAGAGAAACCCCTACAAATGTAAGAATGT 2547
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2548 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 2598

RESULT 5
US-09-949-016-485
; Sequence 485, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 485
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

Alignment Scores:
Pred. No.: 3,548-30 Length: 3839
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-485 (1-3839)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 2408 TGTAAAGAAATGTGGCAAGCAATTATATGCTTCAACCTTACTACTAGATAGAGGATA 2467
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 2468 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 2527
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2528 ACCCTTACTAAGCATAAGACAATTCTACTGGAGAGAAACCCCTACAAATGTAAGAATGT 2587
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2588 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 2638

RESULT 6
US-09-949-016-12227

```

```

; Sequence 12227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 156942
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

Alignment Scores:
Pred. No.: 5,948-28 Length: 156942
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-12227 (1-156942)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 153519 TGTAAAGAAATGTGGCAAGCAATTATATGCTTCAACCTTACTACTAGATAGAGGATA 153578
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 153579 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 153638
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 153639 ACCCTTACTAAGCATAAGACAATTCTACTGGAGAGAAACCCCTACAAATGTAAGAATGT 153698
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 153699 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 153749

RESULT 7
US-09-949-016-15946
; Sequence 15946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15946

```

```
; LENGTH: 156950
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156950)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15946

Alignment Scores:
Pred. No.: 5.94e-28 Length: 156950
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-15946 (1-156950)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 153519 TGTAAAGAAATGTGGCAAGCAATTTATATGGTCTTCAACCCCTAACTAGACATAAGAGGATA 153578
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 153579 CACACTGGAGAGAGCCCTACAAATGTGAAGAAATGTGGCAAGCTTTTAGCGGTTCTCTCA 153638
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 153639 ACCCTTACTAGCATAAGACAATTCATATCTGGAGAGAAACCCCTACAAATGTAAAGAATGT 153698
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 153699 GCAAAAGCTTTTAAAGCACTCTCAGCCCTTCTGCTAAACAATAAATATACAT 153749

RESULT 8
US-09-016-434-1390
; Sequence 1390, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1390:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g468707
US-09-016-434-1390

Alignment Scores:
Pred. No.: 3.68e-30 Length: 3186
Score: 314.00 Matches: 54
Percent Similarity: 80.52% Conservative: 8
Best Local Similarity: 70.13% Mismatches: 15
Query Match: 71.04% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-016-434-1390 (1-3186)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1160 TGTAAAGAAATGTGGCAAGCTTTTCAATCAGAAAGTCAAAACCTCATCAGACACAGAGAACT 1219
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1220 CACACAGAGAGAGAGCCCTTTGTATGTAGAGAGTGTGGAAAAACCTTCAGTGGCAATCC 1279
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1280 AACCTTACTAGCATCAGAAATCCATATTTGGAGAGAAAGCCCTTTTAAATGTAGTGAATGT 1339
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1340 GGAACAGCCCTTTGGCCAGAGAAAGTACCTCATATAAACATCAGAAACATTTCAC 1390

RESULT 9
US-10-104-047-901
; Sequence 901, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 901
; LENGTH: 1835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-901

Alignment Scores:
Pred. No.: 2.31e-30 Length: 1835
Score: 313.00 Matches: 51
Percent Similarity: 81.82% Conservative: 12
Best Local Similarity: 66.23% Mismatches: 14
Query Match: 70.81% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-901 (1-1835)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 693 TGTAAAGGAATGTGGAAAGCCCTTTATCTCTGATTCACATCTTATTCGACATCAGAGTGC 752
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 753 CATACTGGGAGAAACCCCTATATAGTAAAGAAATGTGGAAAGTCCCTTTCGTCGTGGCTCA 812
```

```
QY 41 AsnLeuThrLysHisLysLysLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCys 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 813 GAACCTTACCAGCATCAGAGAGCTCATCTGCTGTAAGAAACCCCTATGAGTGAAGGAATGT 872

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 873 GGAAAGCCCTTACTTGTAGCAGACAGAACTTGTTCGACATCAAAAAGTTTCAC 923

RESULT 10
US-09-949-016-27493
; Sequence 27493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27493
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27493

Alignment Scores:
Pred. No.: 1,61e-30 Length: 601
Score: 309.00 Matches: 51
Percent Similarity: 84.42% Conservative: 14
Best Local Similarity: 66.23% Mismatches: 12
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27493 (1-601)
QY 1 CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeArgHisGlnArgThr 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 TGCAGTGAAGCTGTGGAAAGCCCTTCATTTTCAGTCTTCCCTTAAGAAACACATGAGATCT 82

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 CATACTGGAGAGAGCCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 142

QY 41 AsnLeuThrLysHisLysLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCys 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 CATCTGAATGTGCACAAAAGAACTCACACTGGAGAGAAACCCCTATGAGTGAAGGAATGT 202

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 GGAAGGCTTTCAGTGTTCCTTCATCTCCCTTCAGAAACATGTGAGAACCCAC 253

RESULT 11
US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27493 (1-601)
QY 1 CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeArgHisGlnArgThr 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 TGCAGTGAAGCTGTGGAAAGCCCTTCATTTTCAGTCTTCCCTTAAGAAACACATGAGATCT 82

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 CATACTGGAGAGAGCCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 142

QY 41 AsnLeuThrLysHisLysLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCys 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 CATCTGAATGTGCACAAAAGAACTCACACTGGAGAGAAACCCCTATGAGTGAAGGAATGT 202

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 GGAAGGCTTTCAGTGTTCCTTCATCTCCCTTCAGAAACATGTGAGAACCCAC 253

RESULT 12
US-10-104-047-1500
; Sequence 1500, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1500
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1500

Alignment Scores:
Pred. No.: 6,74e-30 Length: 1694
Score: 309.00 Matches: 52
Percent Similarity: 80.52% Conservative: 10
Best Local Similarity: 67.53% Mismatches: 15
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-1500 (1-1694)
QY 1 CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeArgHisGlnArgThr 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1192 TGTAAAGGAATGTGGCAAGCCCTTTACCAGAGCTTCAAAGCTTGTTCACATCAGAGAATT 1251

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



```
QY      21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db      1813 CATACTGGTGAGAAACCCCTATGAGTGCAGGAATGTGGCAAGCCCTTTAGTAGTGGCTCA 1872
QY      41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db      1873 GCACCTTACTAATCATCAGAGAATTTCACACTGGGTGAGAAACCCCTATGATTGTAAGGAATGT 1932
QY      61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db      1933 GGAAGGCTTTTACTCAGAGCTCACAGCTTCGTCAACATCAGAGAATTCCAC 1983
```

Search completed: January 12, 2006, 03:42:50
Job time : 182 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 01:19:30 ; Search time 614 Seconds
(without alignments)
1037.039 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence:

1 KCDCGKAFIQKSLNRHRT.....KCGKAFGCPNLRHRGRTH 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19597084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US1069861/runat_09012006_144454_20413/app_query.fasta_1.263
-DB=Published Applications NA Main -OFFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	1137	6	US-10-314-669-1
2	442	100.0	1137	8	US-10-669-861-1
3	374	84.6	630	6	US-10-314-669-13
4	374	84.6	630	8	US-10-669-861-13
5	327	74.0	1568	9	US-10-450-763-22733
6	327	74.0	1592	7	US-10-466-164-14
7	327	74.0	2280	6	US-10-108-260A-2425
8	326	73.8	636	6	US-10-314-669-263

9	325	73.5	3300	3	US-09-765-555-18	Sequence 18, Appl
10	321	72.6	330	3	US-09-765-555-64	Sequence 64, Appl
11	321	72.6	3300	3	US-09-765-555-14	Sequence 14, Appl
12	321	72.6	3300	3	US-09-765-555-15	Sequence 15, Appl
13	319	72.2	2175	9	US-10-450-763-18377	Sequence 18377, A
14	319	72.2	2667	6	US-10-108-260A-1226	Sequence 1226, Ap
c 15	318	71.9	505	6	US-10-029-386-2047	Sequence 2047, Ap
c 16	316	71.5	554	4	US-09-925-065A-189236	Sequence 189236,
17	316	71.5	2253	10	US-11-037-295-22	Sequence 22, Appl
18	315	71.3	468	3	US-09-918-995-3107	Sequence 3107, Ap
19	315	71.3	630	6	US-10-314-669-9	Sequence 9, Appl
20	315	71.3	630	8	US-10-669-861-9	Sequence 9, Appl
c 21	315	71.3	632	4	US-09-925-065A-488841	Sequence 488841,
22	315	71.3	1258	6	US-10-029-386-25281	Sequence 25281, A
23	315	71.3	1914	6	US-10-104-047-1464	Sequence 1464, Ap
24	315	71.3	2664	6	US-10-108-260A-511	Sequence 511, App
25	315	71.3	2690	8	US-10-723-860-8039	Sequence 8039, Ap
26	315	71.3	3639	9	US-10-450-763-390	Sequence 390, App
27	315	71.3	3839	6	US-10-172-118-910	Sequence 910, App
28	315	71.3	3839	7	US-10-342-887-910	Sequence 910, App
29	315	71.3	3839	9	US-10-956-157-1072	Sequence 1072, Ap
30	314	71.0	283	6	US-10-029-386-26998	Sequence 26998, A
31	314	71.0	590	6	US-10-029-386-13298	Sequence 13298, A
32	314	71.0	3186	5	US-10-171-581-141	Sequence 141, App
33	314	71.0	3186	6	US-10-305-720-1390	Sequence 1390, Ap
34	314	71.0	3213	5	US-10-098-841-141	Sequence 141, App
35	314	71.0	3382	3	US-09-971-392-83	Sequence 83, Appl
36	313	70.8	1835	6	US-10-104-047-901	Sequence 901, App
37	312	70.6	392	3	US-09-864-761-20124	Sequence 20124, A
38	311	70.4	1945	6	US-10-094-749-1120	Sequence 1120, Ap
39	311	70.4	2662	7	US-10-363-616-66	Sequence 66, Appl
40	310	70.1	630	6	US-10-314-669-259	Sequence 259, App
41	310	70.1	630	8	US-10-669-861-259	Sequence 259, App
42	310	70.1	2873	7	US-10-755-889-369	Sequence 369, App
43	310	70.1	3300	3	US-09-765-555-17	Sequence 17, Appl
44	309	69.9	514	6	US-10-029-386-9811	Sequence 9811, Ap
45	309	69.9	1694	6	US-10-104-047-1500	Sequence 1500, Ap

ALIGNMENTS

RESULT 1

US-10-314-669-1
; Sequence 1, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12779-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

;
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
US-10-314-669-1
Alignment Scores:
Pred. No.: 9,72e-52 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-1 (1-1137)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 97 TGTAAAGATTGCGGGAAGCTTTTCATTTCAGAAAGTCAAAACCTTCATCAGACACCCAGAGAACT 156
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 157 CACACCGGGGAAAAACCGGTACAGGTGTGAAGATGTGGCAAGCTTTTACCCTCAATCCTCA 216
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 217 AACCTTACTAAACATAAGAAATTCATACCGGGGAAAAACCGGTATAATGTAAGCAATGT 276
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 277 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAAGGCATGGAAGGACTCAC 327
RESULT 2
US-10-669-861-1
; Sequence 1, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Lee, Yangsoo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
US-10-669-861-1
Alignment Scores:
Pred. No.: 9,72e-52 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-1 (1-1137)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 97 TGTAAAGATTGCGGGAAGCTTTTCATTTCAGAAAGTCAAAACCTTCATCAGACACCCAGAGAACT 156
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 157 CACACCGGGGAAAAACCGGTACAGGTGTGAAGATGTGGCAAGCTTTTACCCTCAATCCTCA 216
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 217 AACCTTACTAAACATAAGAAATTCATACCGGGGAAAAACCGGTATAATGTAAGCAATGT 276
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 277 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAAGGCATGGAAGGACTCAC 327
RESULT 3
US-10-314-669-13
; Sequence 13, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoo
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; NAME/KEY: CDS
; LOCATION: (1)...(627)
US-10-314-669-13
Alignment Scores:
Pred. No.: 1.67e-42 Length: 630
Score: 374.00 Matches: 61
Percent Similarity: 92.21% Conservative: 10
Best Local Similarity: 79.22% Mismatches: 6
Query Match: 84.62% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-13 (1-630)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 181 TGCAGTGAAGTGTGGCGGAGGCTTCAGCCAGAAAGTCAAAACCTTCATCATACACCCAGAGGACA 240

```



```

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 241 CACACCGGGGAAAAACCGTATGATGTCACGATTGCGGAAAGTCCTTTAGGCAGAGCACC 300
QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 301 CACCTCTACCTCGGACCGGAGGATCCACACCGGGGAAAAACCGTATGATGTCACGATTGCGGAAAGTCCTTTAGGCAGAGCACC 360
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 361 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAGGATGGAAGGACTCAC 411

```

RESULT 4

```

US-10-669-861-13
; Sequence 13, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; NAME/KEY: CDS
; LOCATION: (1)...(627)
US-10-669-861-13

```

```

Alignment Scores:
Pred. No.: 1.67e-42 Length: 630
Score: 374.00 Matches: 61
Percent Similarity: 92.21% Conservative: 10
Best Local Similarity: 79.21% Mismatches: 6
Query Match: 84.62% Indels: 0
DB: 8 Gaps: 0

```

US-10-669-861-2_COPY_33_109 (1-77) x US-10-669-861-13 (1-630)

```

QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
Db 181 TGCAGTGAGTGTGGCGAGGCTTCAGCCAGAGAGTCAAACTCATATACACAGAGGACA 240
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 241 CACACCGGGGAAAAACCGTATGATGTCACGATTGCGGAAAGTCCTTTAGGCAGAGCACC 300
QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 301 CACCTCTACCTCGGACCGGAGGATCCACACCGGGGAAAAACCGTATGATGTCACGATTGCGGAAAGTCCTTTAGGCAGAGCACC 360
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 361 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAGGATGGAAGGACTCAC 411

```

```

RESULT 5
US-10-450-763-22733
; Sequence 22733, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22733
; LENGTH: 1668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (495)..(584)
; OTHER INFORMATION: 23% homologous to Homo sapiens zinc finger protein
; OTHER INFORMATION: ZNF136, accession number U09367, Smith-Waterman Score=43.
US-10-450-763-22733

```

```

Alignment Scores:
Pred. No.: 2.47e-35 Length: 1668
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 9 Gaps: 0

```

US-10-669-861-2_COPY_33_109 (1-77) x US-10-450-763-22733 (1-1668)

```

QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
Db 595 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATGTCCTCACACCTTCAAAATGATGATGAGGACT 654
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 655 CACACTGGAGAGAAACCTCTATGATGTAAGCAGTGTGGGAAAGTCCTTTGGATGTCCTCG 714
QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 715 CGACTTCAATGTCATGGAGGACTCACACTGGAGAGAAACCTGATTAATGTAAGCAATGT 774
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 775 GGGAAAGCTTTTGGATGTCCTCAAACTTCGAGGATGGAAGGACTCAC 825

```

RESULT 6

```

US-10-466-164-14
; Sequence 14, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFPO, Abel;
; APPLICANT: MARWAHA, Rakeeh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN, Jr., Edward H.;
; APPLICANT: PERALTA, Careytha A.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT

```

```

; CURRENT APPLICATION NUMBER: US/10/466,164
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 2192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 Li.1175131.1:2001JAN12
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 64, 81
; OTHER INFORMATION: a, t, c, g, or other
US-10-466-164-14

Alignment Scores:
Pred. No.: 3,54e-35 Length: 2192
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-466-164-14 (1-2192)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1845 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1904

Qy 21 HistThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1905 CACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGTGGGAAAGTCTTTGGATGTCCTCG 1964

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1965 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGT 2024

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2025 GGGAAAGCTTTTGGATGTCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 2075

RESULT 7
US-10-108-260A-2425
; Sequence 2425, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2425
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2425

Alignment Scores:
Pred. No.: 3,73e-35 Length: 2280
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-108-260A-2425 (1-2280)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1861 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1920

Qy 21 HistThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1921 CACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGTGGGAAAGTCTTTGGATGTCCTCG 1980

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1981 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGT 2040

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2041 GGGAAAGCTTTTGGATGTCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 2091

RESULT 8
US-10-314-669-263
; Sequence 263, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(633)
US-10-314-669-263

Alignment Scores:
Pred. No.: 9,61e-36 Length: 636

```

```

Score: 326.00 Matches: 55
Percent Similarity: 85.71% Conservative: 11
Best Local Similarity: 71.43% Mismatches: 11
Query Match: 73.76% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-263 (1-636)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 97 TGTAAAGATTGCGGGAAGCTTCATTCAGAACTCAACTCATCAGACACAGAACT 156
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 157 CACACCGGGGAAAAACCGTATAGTCCCTGATGTTGGGAAGAGTTTATGTCAGAGTTC 216
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 217 AGCTCATTCGCCACCGACCGGACACACACCGGGGAAAAACCGTATGAGTGTCCAGATTGC 276
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 277 GGAAAGTCTTTAGGCAGACGACCCACTCTCGGACCGGAGGATCCAC 327

RESULT 9
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZFPp3
US-09-765-555-18
; OTHER INFORMATION: Parial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPp3
US-09-765-555-18

Alignment Scores:
Pred. No.: 1,166-34 Length: 3300
Score: 325.00 Matches: 54
Percent Similarity: 84.4% Conservative: 11
Best Local Similarity: 70.13% Mismatches: 12
Query Match: 73.53% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-765-555-18 (1-3300)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 2758 TGTCCGAATGTGGTAAGCTTCAGCCAGACAGCTCCTGTCGCGCCACAGCGTACC 2817
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 2818 CACACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTAGCCAGTCCAGC 2877
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 2878 AACCTGTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGT 2937
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77

```

```

Db 2938 GGCAAGTCTTTAGCCAGACCTGTCGCGCCACCAAGTACTCAC 2988

RESULT 10
US-09-765-555-64
; Sequence 64, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZFP2b
US-09-765-555-64

Alignment Scores:
Pred. No.: 2,056-35 Length: 330
Score: 321.00 Matches: 53
Percent Similarity: 83.12% Conservative: 11
Best Local Similarity: 68.83% Mismatches: 13
Query Match: 72.62% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-765-555-64 (1-330)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 55 TGTCCGAATGTGGTAAGCTTCCTCTCAGAGCTCTCACCTGTCGCGCCACAGCGTACC 114
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 115 CACACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTAGCCAGTCCAGC 174
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 175 AACCTGTGCGCCATCAACGCACTCATCTGTCGCGGAGAGCCATACAAATGTCCAGAAATGT 234
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 235 GGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCGCGGACCAACGACTACTCAC 285

RESULT 11
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300

```



```

: TITLE OF INVENTION: NO. US2004000556A1el full length cDNA
:
: FILE REFERENCE: H1-A0106
: CURRENT APPLICATION NUMBER: US/10/108,260A
: CURRENT FILING DATE: 2002-03-27
: NUMBER OF SEQ ID NOS: 5458
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1226
: LENGTH: 2667
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-108-260A-1226

```

Alignment Scores:		
Pred. No.:	6.12e-34	2667
Score:	319.00	54
Percent Similarity:	84.4%	Conservative: 11
Best Local Similarity:	70.1%	Mismatches: 12
Query Match:	72.1%	Indels: 0
DB:	6	Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-108-260A-1226 (1-2667)

QY	1	CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr	20
DB	1412	TGTGAAGAAGTGTGCAAAAGCTTTTAAACCAAGTCCTCATCCCTTACTTAACATAAGAAAATT	1471
QY	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
DB	1472	CATCTCGAGAGAAACCCCTCAAAATGTGAAGAATGTGCAAAGCTTTTAAACCAAGTCCTCT	1531
QY	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
DB	1532	TCCCTTACTTAACATAAGAAAATTCATCTGGAGAGAAAACCTTACAATGTGAAGAATGT	1591
QY	61	GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis	77
DB	1592	GGCAAGCTTTTAAACCAAGTCCTCAACCTTATTAAAATGAAGAAAATTCAT	1642

```

RESULT 15
US-10-029-386-2047/c
; Sequence 2047, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2047
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
; OTHER INFORMATION: EST_HUMAN HIT: AW236498.1, EVALUAE 0.00e+00
; OTHER INFORMATION: NT HIT: G114760087, EVALUAE 5.00e-35
; OTHER INFORMATION: SWISSPROT HIT: P17024, EVALUAE 6.00e-44
US-10-029-386-2047

```

Alignment Scores:	
Pred. No.:	9.49e-35
Length:	505

Score:	318.00	Matches:	55
Percent Similarity:	81.58%	Conservative:	7
Best Local Similarity:	72.37%	Mismatches:	14
Query Match:	71.95%	Indels:	0
DB:	6	Gaps:	0
US-10-669-861-2_COPY_33_109 (1-77) x US-10-029-386-2047 (1-505)			
QY	2	LYSASP	CYSGLYLVEALAPHELLEGLNLYSSERASNLEULLEIARGHISGLINAA
DB	503	AAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAA	
QY	22	THRGLYGLULYSPROTYRLYSCYSGLUGLUCYSGLYLYSALAPHE	THRGLINS
DB	443	ACTGGAGAGAACCCCTATGAATGAAGCAGTGTGGGAAGTCTTTTGGATGTG	
QY	42	LEUTHRLYSHISLYSLEHISLTHRGLYGLULYSPROTYRLYSCYSLYSGL	
DB	383	CTTCAAAATGCATGGGAAGGACTCACACTGGAGAGAAACCGTATAAATCTAAGC	
QY	62	LYSALAPHEGLYCYP	PROSERASNLEUARGHISGLYARGTHRHS
DB	323	AAAGCTTTTGGATGTCCCTCAAACCTTCAAGGCATGGGAAGGACTCAC	276

Search completed: January 12, 2006, 03:39:45
Job time : 616 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 02:31:44 ; Search time 226 Seconds
(without alignments)
275.751 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCKAFQKSNLRHQRTH.....KQCKAFGCPNLRHRGRTH 77

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144454_20431/app_query.fasta_1.263
-DB=PublishedApplications_NA_New -Qfmt=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10669861@cgn 1.122 @runat_09012006_144454_20431
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	302	68.3	982	6	US-10-750-185-60532 Sequence 60532, A
C 2	302	68.3	982	6	US-10-750-623-54987 Sequence 54987, A
C 3	302	68.3	2196	7	US-11-000-688-253 Sequence 253, App
4	301	68.1	2124	7	US-11-136-527-713 Sequence 713, App
5	299	67.6	600	7	US-11-136-527-4391 Sequence 4391, App
6	299	67.6	1940	7	US-11-136-527-295 Sequence 295, App
7	298	67.4	600	7	US-11-136-527-6532 Sequence 6532, App
8	298	67.4	1578	7	US-11-136-527-2436 Sequence 2436, App

ALIGNMENTS

RESULT 1

US-10-750-185-60532/c
; Sequence 60532, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 60532
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-60532

Alignment Scores:
Pred. No.: 1.15e-28 Length: 982
Score: 302.00 Matches: 50
Percent Similarity: 81.82% Conservative: 13
Best Local Similarity: 64.94% Mismatches: 14
Query Match: 68.33% Indels: 0

Sequence 54987, A
Sequence 54987, A
Sequence 44166, A
Sequence 44166, A
Sequence 6854, App
Sequence 2758, App
Sequence 41858, A
Sequence 41858, A
Sequence 3886, App
Sequence 601, App
Sequence 129, App
Sequence 1179, App
Sequence 1046, App
Sequence 5142, App
Sequence 710, App
Sequence 28, Appl
Sequence 119, App
Sequence 55, Appl
Sequence 42613, A
Sequence 42613, A
Sequence 30010, A
Sequence 30010, A
Sequence 49223, A
Sequence 49223, A
Sequence 928, App
Sequence 61337, A
Sequence 61337, A
Sequence 58156, A
Sequence 58156, A
Sequence 55, Appl
Sequence 138, App
Sequence 1136, App
Sequence 2585, App
Sequence 6227, App
Sequence 49213, A
Sequence 49213, A
Sequence 5181, App

```
DB:
US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-60532 (1-982)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 477 TGTAAAGAGTGTAAACAAAGCCTTTATTACGCGCTCACATCTTACTATAACATCAGAAAAATT 418
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 417 CATACTGGAGAGACCTTTATAATGTAAAGAGTGTAAACAAAGCCTTTATTACGCGCGCA 358
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 357 AATCTTACTAAACATCAGCGAATTCATACGGGAGAGACCTTTATAATGTAAAGAGTGT 298
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 297 AACAAAGCCTTTAATCAACGCTCAAGTCTTAACCTATCATCAGCGAATTCAT 247
RESULT 2
US-10-750-623-60532/c
; Sequence 60532, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60532
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Bovine 19866881803796
US-10-750-623-60532
Alignment Scores:
Pred. No.: 1.15e-28 Length: 982
Score: 302.00 Matches: 50
Percent Similarity: 81.82% Conservative: 13
Best Local Similarity: 64.94% Mismatches: 14
Query Match: 68.33% Indels: 0
DB:
US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-60532 (1-982)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 477 TGTAAAGAGTGTAAACAAAGCCTTTATTACGCGCTCACATCTTACTATAACATCAGAAAAATT 418
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 417 CATACTGGAGAGACCTTTATAATGTAAAGAGTGTAAACAAAGCCTTTATTACGCGCGCA 358
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 357 AATCTTACTAAACATCAGCGAATTCATACGGGAGAGACCTTTATAATGTAAAGAGTGT 298
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 297 AACAAAGCCTTTAATCAACGCTCAAGTCTTAACCTATCATCAGCGAATTCAT 247
RESULT 3
US-10-669-861-2_COPY_33_109 (1-77) x US-11-000-688-253 (1-2196)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 1476 TGTAAAGACTGTGGGAAGGCTTTTATTGTGGTTCAAGCCTTCATTCAGCATAAAGAATT 1535
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1536 CACACAGGTGGAACCTTATGATGTCAAGAATGTGGGAAGGCTTTACTCGAGTCAAT 1595
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 1596 TACCTTACTCAGCATCAGAAGATCCACACCGGTGAGAAGCCTCAGAAATGTAAAGAGTGT 1655
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1656 GGGAAAGCCTTTCTCGTGGGTTTCGAGGCTCGTTAAGCACGAGAGATACAT 1706
RESULT 4
US-11-136-527-713
; Sequence 713, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 713
; LENGTH: 2124
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-713
US-11-000-688-253
; Sequence 253, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 253
; LENGTH: 2196
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc feature
; LOCATION: (1)-(2196)
; OTHER INFORMATION: zinc finger protein 463 (ZNF463) gene.
US-11-000-688-253
Alignment Scores:
Pred. No.: 3.22e-28 Length: 2196
Score: 302.00 Matches: 52
Percent Similarity: 85.71% Conservative: 14
Best Local Similarity: 67.53% Mismatches: 11
Query Match: 68.33% Indels: 0
DB:
```


Best Local Similarity: 67.53% Mismatches: 17
 Query Match: 67.42% Indels: 0
 DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6532 (1-600)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 Db 19 TGTAAAGGGTGTGGAGGCGCTTATTAGAAATCAAGCCTCATCAGACACAGAGGAGC 78
 Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 79 CATACTGGAGAGAAACCGTACACATGTAAGGAATGTGGAAAGCCTTCAGCGGCAATCA 138
 Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
 Db 139 AACCTGACTGAGCATGAGAAATTCATATTGGAGAGAAACCCCTATAAATGTAACGAGTGT 198
 Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 199 GGAACAATTTTCAGGCAGACAGTACCTCATCAAAATCATCAACATTCAC 249

RESULT 8

US-11-136-527-2436
 ; Sequence 2436, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2436
 ; LENGTH: 1578
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-2436

Alignment Scores:
 Pred. No.: 6,866-28 Length: 1578
 Score: 298.00 Matches: 52
 Percent Similarity: 77.92% Conservative: 8
 Best Local Similarity: 67.53% Mismatches: 17
 Query Match: 67.42% Indels: 0
 DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-2436 (1-1578)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 Db 997 TGTAAAGGGTGTGGAGGCGCTTATTAGAAATCAAGCCTCATCAGACACAGAGGAGC 1056
 Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 1057 CATACTGGAGAGAAACCGTACACATGTAAGGAATGTGGAAAGCCTTCAGCGGCAATCA 1116
 Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
 Db 1117 AACCTGACTGAGCATGAGAAATTCATATTGGAGAGAAACCCCTATAAATGTAACGAGTGT 1176
 Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 1177 GGAACAATTTTCAGGCAGACAGTACCTCATCAAAATCATCAACATTCAC 1227

RESULT 9

US-10-750-185-54987/c
 ; Sequence 54987, Application US/10750185
 ; Publication No. US20050260603A1

GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: KERR, Richard
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: HOLM, Tom
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: FANTIN, Dennis
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MM1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 54987
 ; LENGTH: 984
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866880605923
 US-10-750-185-54987

Alignment Scores:
 Pred. No.: 6,786-28 Length: 984
 Score: 296.00 Matches: 51
 Percent Similarity: 79.22% Conservative: 10
 Best Local Similarity: 66.23% Mismatches: 16
 Query Match: 66.97% Indels: 0
 DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-54987 (1-984)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 Db 609 TGCACCTAAGTGTGGAAAGGCTTTCAGTAGGAATCCAGGCTTGTGAGCATCAGAGAACT 550
 Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 549 CATACAGGTGAAAAACCCCTAGAGTGGGTGAATGTGGAAAGCTTTCAGTGAATAATTA 490
 Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
 Db 489 AGCCTCACTAATCACCAGAGAAATTCATACAGGAGAAACCGTATGTGTGCGAGTGAGTGT 430
 Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 429 GGGAAAGCCTTTTGTGCAAGTTCATCTCATCATCATCAGAGGACACAC 379

RESULT 10

US-10-750-623-54987/c
 ; Sequence 54987, Application US/10750623
 ; Publication No. US20050287531A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: KERR, Richard
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: HOLM, Tom
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: FANTIN, Dennis
 ; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MM1100-1
 ; CURRENT APPLICATION NUMBER: US/10/750,623
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 54987
 ; LENGTH: 984
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866880605923
 US-10-750-623-54987

Alignment Scores:
Pred. No.: 6,78e-28 Length: 984
Score: 296.00 Matches: 51
Percent Similarity: 79.22% Conservatives: 10
Best Local Similarity: 66.23% Mismatches: 16
Query Match: 66.97% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-54987 (1-984)

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 609 TGCACTAAGTGGGAGGCTTCAGTAGAATCCAGCTTGTCCGACATCAGAGACT 550

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 549 CATACAGGTGAAACCCATCAGGTGCGTGAATGGGAAAGCTTCAGTGAATAA 490

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 489 AGCTCACTAATCACCAGAGAATTCATACAGGAGAAACCGTATGTGTGAGTGAGTGT 430

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 429 GCGAAGGCTTTGTGTCAGAGTCACATCTCATATCAGATCAGAGACAC 379

RESULT 11
US-10-750-185-44166/c
; Sequence 44166, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-185-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservatives: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-750-185-44166

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 3428 TGTAAGAAATGAGAAAGCTTCAGCAAAATGTACATCTATTCAACATCAAGAAT 3369

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 3368 CATACGTGAGAAACCATATAAATGTAAGAGTGTAGAAAGCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 3308 CACCTTGCTCAGCATCAGAGAATTCATACCTGGGAGAGGCTTTATAAATGTAAGGAATGT 3249

RESULT 12
US-10-750-623-44166/c
; Sequence 44166, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-623-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservatives: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-44166 (1-4305)

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 3428 TGTAAGAAATGAGAAAGCTTCAGCAAAATGTACATCTATTCAACATCAAGAAT 3369

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 3368 CATACGTGAGAAACCATATAAATGTAAGAGTGTAGAAAGCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 3308 CACCTTGCTCAGCATCAGAGAATTCATACCTGGGAGAGGCTTTATAAATGTAAGGAATGT 3249

RESULT 13
US-11-136-527-6854
; Sequence 6854, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2

Alignment Scores:
Pred. No.: 6,78e-28 Length: 984
Score: 296.00 Matches: 51
Percent Similarity: 79.22% Conservatives: 10
Best Local Similarity: 66.23% Mismatches: 16
Query Match: 66.97% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-54987 (1-984)

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 609 TGCACTAAGTGGGAGGCTTCAGTAGAAATCCAGCTTGTCCGACATCAGAGACT 550

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 549 CATACAGGTGAAACCCATCAGGTGCGTGAATGGGAAAGCTTCAGTGAATAA 490

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 489 AGCTCACTAATCACCAGAGATTATACAGGAGAAACCGTATGTGTGAGTGAGTGT 430

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 429 GCGAAGGCTTTGTGAGAGTCACATCTCATATCAGATCAGAGACAC 379

RESULT 11
US-10-750-185-44166/c
; Sequence 44166, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-185-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservatives: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-750-185-44166

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 3428 TGTAAGAAATGAGAAAGCTTCAGCAAAATGTACATCTATTCAACATCAAGAATT 3369

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 3368 CATACTGGAGAAACCATATAAATGTAAGAGTGTAGAAAGCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 3308 CACCTTGCTCAGCATCAGAGAAATTCATACCTGGGAGAGAGCTTTATAAATGTAAGGAATGT 3249

RESULT 12
US-10-750-623-44166/c
; Sequence 44166, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-623-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservatives: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-44166 (1-4305)

QY 1 CysLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 20
Db 3428 TGTAAGAAATGAGAAAGCTTCAGCAAAATGTACATCTATTCAACATCAAGAATT 3369

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 3368 CATACTGGAGAAACCATATAAATGTAAGAGTGTAGAAAGCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 3308 CACCTTGCTCAGCATCAGAGAAATTCATACCTGGGAGAGAGCTTTATAAATGTAAGGAATGT 3249

RESULT 13
US-11-136-527-6854
; Sequence 6854, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2

```
; SEQ ID NO 6854
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6854

Alignment Scores:
Pred. No.: 2, 57e-27 Length: 1400
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6854 (1-1400)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1047 TGTGAAGAATGTGGAAAGCCTTTGGCTGTTAAATCAAACTTTATCGGCATCAGAGGATC 1106
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1107 CATACTGGTGAGAAACCCCTATCAGTGTAAATCAGTGTGGCAAGGCCTTCAGCCAGTATTCA 1166
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1167 TTCCTTAACGGACACAGCGAATCCATCTGAGAGAAAGCTCTACAGTGTATGGAGTGT 1226
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1227 GGGAAAGCCTATAGTTACAGGTCAAACTCTGTAGACACAAAGAGTCCAC 1277

RESULT 14
US-11-136-527-2758
; Sequence 2758, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2758
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2758

Alignment Scores:
Pred. No.: 3, 25e-27 Length: 1683
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-2758 (1-1683)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1330 TGTGAAGAATGTGGAAAGCCTTTGGCTGTTAAATCAAACTTTATCGGCATCAGAGGATC 1389
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1390 CATACTGGTGAGAAACCCCTATCAGTGTAAATCAGTGTGGCAAGGCCTTCAGCCAGTATTCA 1449
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1449 TGTGAAGAATGTGGAAAGCCTTTGGCTGTTAAATCAAACTTTATCGGCATCAGAGGATC 1508

; SEQ ID NO 6854
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6854

Alignment Scores:
Pred. No.: 2, 57e-27 Length: 1400
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6854 (1-1400)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1047 TGTGAAGAATGTGGAAAGCCTTTGGCTGTTAAATCAAACTTTATCGGCATCAGAGGATC 1106
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1107 CATACTGGTGAGAAACCCCTATCAGTGTAAATCAGTGTGGCAAGGCCTTCAGCCAGTATTCA 1166
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1167 TTCCTTAACGGACACAGCGAATCCATCTGAGAGAAAGCTCTACAGTGTATGGAGTGT 1226
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1227 GGGAAAGCCTATAGTTACAGGTCAAACTCTGTAGACACAAAGAGTCCAC 1277

RESULT 15
US-10-750-185-41858
; Sequence 41858, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41858
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Bovine 19866881116464
US-10-750-185-41858

Alignment Scores:
Pred. No.: 1, 05e-27 Length: 438
Score: 291.00 Matches: 51
Percent Similarity: 72.73% Conservative: 5
Best Local Similarity: 66.23% Mismatches: 21
Query Match: 65.84% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-41858 (1-438)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 38 TGTGATGACTGTGCAAGGCCTTTTCGTTTAAAGTCATCTCTTTTAACTCATCAGACAGTT 97
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 98 CATACTGGAGAGAAACCTTACAAATGTGATGAGTGTGGCAAGGCCTTTTCGTCGAAAGTCA 157
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 158 ACCCTTTTATCTCATCAGACAATACATCTGGAGAGAAACCTTACAAATGTAAATGAGTGT 217
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 218 GCGAAGGTCTTTACTCATCAGCTCAAAATCTCAGGAGACATCAGAAAAATTCAT 268

Search completed: January 12, 2006, 03:46:18
Job time : 228 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 00:39:24 ; Search time 3025 Seconds

(without alignments)
1190.944 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCKAFQKSNLRHRT.....KCGKAFGCPNLRHRGTH 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144453_20382/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOCPCL=0 -LOCPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOCALIGN=200 -NORM SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10669861@CNC 1 1 5315 @runat_09012006_144453_20382 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	74.4	458	1 AJ673279	AW236498
2	327	74.0	518	1 AW236498	CR849534
3	326	73.8	889	7 CR849534	CZ921322
4	326	73.8	922	10 CZ921322	BI654450
5	316	71.5	923	3 BI654450	BG245794
6	316	71.5	990	2 BG245794	BG262051
7	316	71.5	1015	2 BG262051	

CF145979	UI-HF-CB0	536	71.3	315	6	CF145979
CB409262	NISC.nc04	632	71.3	315	6	CB409262
AU123664	AU123664	827	71.3	315	1	AU123664
BQ221286	AGENCOURT	875	71.3	315	3	BQ221286
BM904772	AGENCOURT	1066	71.3	315	3	BM904772
DQ030160	Homo sapi	1362	71.3	315	11	DQ030160
BX483171	DKFZp686B	441	71.0	314	5	BX483171
DQ045809	Pan trogl	584	71.0	314	11	DQ045809
BQ306591	MSO-BT300	592	71.0	314	3	BQ306591
AQ356926	CITBI-E1	597	71.0	314	9	AQ356926
AW970087	EST382169	675	71.0	314	1	AW970087
BM973066	UI-CF-ECL	691	71.0	314	3	BM973066
CA509386	UI-R-FSO-	724	71.0	314	6	CA509386
CB325611	UI-R-D20-	765	71.0	314	6	CB325611
BU569663	AGENCOURT	780	71.0	314	5	BU569663
CD623771	56084457J	807	71.0	314	6	CD623771
BG482615	602502762	830	71.0	314	2	BG482615
CJ011854	CJ011854	833	71.0	314	7	CJ011854
CJ021276	CJ021276	843	71.0	314	7	CJ021276
CJ038720	CJ038720	847	71.0	314	7	CJ038720
BG619358	602619493	855	71.0	314	2	BG619358
CR981016	CR981016	865	71.0	314	7	CR981016
BP438052	BP438052	874	71.0	314	3	BP438052
CJ012987	CJ012987	875	71.0	314	7	CJ012987
DQ045808	Homo sapi	879	71.0	314	11	DQ045808
BP463090	BP463090	884	71.0	314	3	BP463090
BP439903	BP439903	887	71.0	314	3	BP439903
BP433982	BP433982	890	71.0	314	3	BP433982
BM541411	AGENCOURT	1102	71.0	314	3	BM541411
DQ032737	Homo sapi	1580	71.0	314	11	DQ032737
DQ036016	Homo sapi	1836	71.0	314	11	DQ036016
BC005154	Homo sapi	2197	71.0	314	4	BC005154
CR859112	Pongo PV9	3126	71.0	314	4	CR859112
CR857109	Pongo PV9	3239	71.0	314	4	CR857109
CB809285	AMGNNUC:N	382	70.8	313	6	CB809285
BB834711	BB834711	461	70.8	313	2	BB834711
AI788680	UK52F01.X	527	70.8	313	1	AI788680
AZ998278	2M0285G08	540	70.8	313	9	AZ998278

ALIGNMENTS

RESULT 1

AJ673279

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 458)

Anderson, S.I., Finlayson, H.A. and Archibald, A.L.

Development of cDNA and EST resources for studying reproduction and

embryo development in pigs and cattle

Unpublished (2004)

Contact: Anderson SI

Genomics and Bioinformatics

Roslin Institute

Roslin, Midlothian, EH25 9PS, UNITED KINGDOM

Single pass sequencing. Bases called and trimmed with phred

v0.020425.c. Vector identified by cross match with the -m9score 20

and -minmatch 12 options. Vector: pBluescriptII(SK+) R. Site 1:

ECORI R. Site 2: NotI 5' Seq primer M13F Description: Normalised

library constructed from Bovine Uterus tissue. Clones available

from UK Centre for Functional Genomics in Farm Animals, Roslin

Institute, Roslin, Midlothian, UK, EH25 9PS, www.arkgenomics.org.

Location/Qualifiers

1. 458

/organism="Bos taurus"

```
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="KN224-011_117"
/tissue_type="uterus"
/clone_lib="KN224"
/notes="Vector: pBluescriptII(SK+); Site 1: EcoRI; Site 2:
NotI; Single pass sequencing. Normalised library
constructed from Bovine Uterus tissue."

ORIGIN

Alignment Scores:
Pred. No.: 4.46e-28 Length: 458
Score: 329.00 Matches: 56
Percent Similarity: 85.71% Conservative: 10
Best Local Similarity: 72.73% Mismatches: 11
Query Match: 74.43% Indels: 0
DB: 1 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AJ673279 (1-458)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 27 TGCACAGATTGCGCAAGCCCTTCAGTCACGCCCTTATCCGTCATCAGAGAAC 86
Qy 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 87 CACACTGGAGAGAACCTTACAGTGTAGAGACTGTGGAGAAAGCCTTCAGCGAGACTCA 146
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysGlyLysProTyrLysCysLysGlnCys 60
Db 147 TCTCTTCAAGAGCATCAGAAACTCACACTGGAGAGAAACCTTATAAGTGTAAAGGAATGT 206
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 207 GGAAGGGCTTTCAGCCAGAGTTCATCCCTTTCTCAACATCAGAAACTCAT 257

RESULT 2
AW236498/c AW236498 518 bp mRNA linear EST 13-DEC-1999
LOCUS xm7405.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2687312 3'
DEFINITION similar to SW:Z136 HUMAN P52737 ZINC FINGER PROTEIN 136. ;contains
Alu repetitive element; mRNA sequence.
ACCESSION AW236498.1 GI:6568887
VERSION AW236498
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 518)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 400.
Location/Qualifiers
1..518
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
1..518
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:2687312"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clones
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

ORIGIN

Alignment Scores:

Pred. No.:	8.87e-28	Length:	518
Score:	327.00	Matches:	56
Percent Similarity:	81.82%	Conservative:	7
Best Local Similarity:	72.73%	Mismatches:	14
Query Match:	73.98%	Indels:	0
DB:	1	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x AW236498 (1-518)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20

Db 449 TGTAAAGCAGTGTGGGAAGCCTTTGGATCTGCCCTCACACCTTCAATGATCGATGGAGGACT 390

Qy 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40

Db 389 CACACTGGAGAGAACCTTATGAATGTAAAGCAGTGTGGGAAGTCTTTTGGATGTGCCTCG 330

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysGlyLysProTyrLysCysLysGlnCys 60

Db 329 CGACTTCAATGTCATGGAAGGACTCACACTGGAGAGAAACCGTATAATGTAAGCAATGT 270

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77

Db 269 GGAAGGCTTTTGGATGTCCCTCAACCTTCGAAGGCATGGAAGGACTCAC 219

RESULT 3

CR849534 889 bp mRNA linear EST 23-MAR-2005

LOCUS CR849534

DEFINITION Normalized and Subtracted bovine endometrium tissues (bcbp) Bos taurus cDNA clone bcbp0003.f.09 5', mRNA sequence.

ACCESSION CR849534.2 GI:61725063

VERSION CR849534

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 889)

Renard, J. P., Lewin, H. A., Yang, J., Hernandez, A., Sandra, O., Everts, R. E. and Hue, I.

Unpublished (1997)

Endometrium ESTs (bcbp)

On Oct 21, 2004 this sequence version replaced gi:54564405.

Contact: Renard JP

Biologie du Developpement et Reproduction

INRA

Domaine de Vilvert 78350 Jouy en Josas, FRANCE

Email: renard@jouy.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us at sigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence. AGENAE Resource centre. Francois PIUMI.

sequence. AGENAE Resource centre. Francois PIUMI.

France. Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (UREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73

Plate: bcbp0003c row: f column: 09.

Location/Qualifiers

FEATURES

```

source
1. 889
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="bcbp003.f.09"
/tissue_type="caruncular and inter-caruncular regions from
embryo mortality, day 20 cyclic cows"
/clone_lib="Normalized and subtracted bovine endometrium
tissues (bcbp)"
/notes="Vector: pBluescript II SK(+); Funding for
construction and subtraction of the cDNA library was
provided to J. P. Renard (BDR, INRA). Funding for EST
sequencing was provided by a Genanimal contract from
FNS-Ministere de la Recherche (INRA, Num: A02085). The
cDNA library was constructed, normalized and subtracted by
Dr. A. Hernandez (W.M. Keck Center, University of Illinois
Urbana-Champaign) and Dr O. Sandra (BDR, INRA) as
described in Bonaldo et al.(1996), Genome Research 6,
791-806. The BCBP library is a normalized library of
endometrium (caruncular and inter-caruncular regions)
subtracted with previously sequenced clones from bovine
placenta (BP), bovine embryo (BE) and bovine spleen (BS)
libraries produced in the laboratory of Dr. H.A. Lewin
(University of Illinois Urbana-Champaign). The double
stranded cDNA was size selected (more than 450 bp). Size
selected cDNA was ligated to EcoRI adaptors, digested with
NotI then directionally cloned into EcoRI-NotI digested
pBluescript II SK(+) phagemid vector. Insert size of the
bcbp was between 600-3500 bp."

ORIGIN
Alignment Scores:
Pred. No.: 2,27e-27 Length: 889
Score: 326.00 Matches: 55
Percent Similarity: 84.42% Conservative: 10
Best Local Similarity: 71.43% Mismatches: 12
Query Match: 73.76% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x CR849534 (1-889)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
Db 66 TGTAAAGATGTGGAAAGCCCTTTATTCAGAGCTCAGAACTATTCAACATCAGAGAATC 125
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 126 CATACAGGTGAAAAACCATATGAATGATGAGTGTGGAAGGCTTTTAAATAAGGCTCA 185
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 186 AACCTTACTCGTCATCAAGAATTACACTGGTGAGAAACCTTATGACTGTAAAGATGT 245
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 246 GGAAGTCTCTTGGTAGTCGCTCGACCTCATCCGCCATGAAGGAATCAT 296

RESULT 4
C2921322
LOCUS
DEFINITION
survey sequence.
ACCESSION
C2921322
VERSION
C2921322.1 GI:71943436
KEYWORDS
GSS.
SOURCE
Ovis aries (sheep)
ORGANISM
Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Caprinae; Ovis.
REFERENCE
1 (bases 1 to 922)
AUTHORS
Kirkness,E.

C2921322
LOCUS
DEFINITION
922 bp DNA linear GSS 08-AUG-2005
survey sequence.
ACCESSION
C2921322
VERSION
C2921322.1 GI:71943436
KEYWORDS
GSS.
SOURCE
Ovis aries (sheep)
ORGANISM
Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Caprinae; Ovis.
REFERENCE
1 (bases 1 to 922)
AUTHORS
Kirkness,E.

```

```

TITLE
JOURNAL
COMMENT
Ovine BAC Ends
Unpublished (2005)
Other_GSSs: 1098415567615
Contact: Ewen Kirkness
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-795-7536
Email: ekirknes@tigr.org
Sequences generated at the J. Craig Venter Institute Joint
Technology Center (JCVI/JTC; http://www.venterlinstitute.org/)
Insert Length: 184000 Std Error: 0.00 row: C column: 22
Seq primer: FORWARD
Class: BAC ends
High quality sequence start: 78
High quality sequence stop: 892.
Location/Qualifiers
1. 922
/organism="Ovis aries"
/mol_type="genomic DNA"
/db_xref="taxon:9940"
/clone="CH243-350C22"
/clone_lib="CH243"
/notes="Vector: PTARBAC2.1"

ORIGIN
Alignment Scores:
Pred. No.: 2,37e-27 Length: 922
Score: 326.00 Matches: 55
Percent Similarity: 84.42% Conservative: 10
Best Local Similarity: 71.43% Mismatches: 12
Query Match: 73.76% Indels: 0
DB: 10 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x C2921322 (1-922)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
Db 302 TGTAAAGATGTGGAAAGCCCTTTATTCAGAGCTCAGAACTATTCAACATCAGAGAATC 361
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 362 CATACAGGTGAAAAACCATATGAATGATGAGTGTGGAAGGCTTTTAAATAAGGCTCA 421
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlnCys 60
Db 422 AACCTTACTCGTCATCAAGAATTACACTGGTGAGAAACCTTATGACTGTAAAGATGT 481
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 482 GGAAGTCTCTTGGTAGTCGCTCGACCTCATCCGCCATGAAGGAATCAT 532

RESULT 5
BI654450
LOCUS
DEFINITION
603281023F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5325509 5',
mRNA sequence.
ACCESSION
BI654450
VERSION
BI654450.1 GI:15568686
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 923)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
CONTACT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.

```


CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM11825 row: n column: 06
 High quality sequence stop: 699.
 Location/Qualifiers
 1..923
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NNRI"
 /db_xref="taxon:10090"
 /clone="IMAGE:5325509"
 /tissue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam4"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

FEATURES

source

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10331 row: i column: 22
 High quality sequence stop: 701.
 Location/Qualifiers
 1..990
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="INAGR:4487133"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="3 months, virgin"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam1"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 3,456-26 Length: 923
 Score: 316.00 Matches: 53
 Percent Similarity: 81.82% Conservative: 10
 Best Local Similarity: 68.83% Mismatches: 14
 Query Match: 71.49% Indels: 0
 DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x B1654450 (1-923)

Qy 1 CysLysAspCysGlyLysAlaPheileGlnLysSerAsnLeuileArgHisGlnArgThr 20
 Db 14 TGTGACAGTGTGATAGGCAATTTGCACAAACAGTCATCTTAAACACATAAAGAAGACT 73
 Qy 21 HistHrGlyGluLysProTyrLysCysGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 74 CATACTGGAGAGAAACCCATGATGTGAGCAATGTGCAAAAGCCTTTGCAAGTAATAGT 133
 Qy 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 134 AATCTGCAAGTGCATAAAAAATGCACACTGGAGAGAAACCCATGATGTAAAGCAGTGT 193
 Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 194 GGTAAGCCTTTGGGTATCAAAAGTGGTCTGCAAAAGCATAAACGACACAT 244

RESULT 6
 BG245794
 LOCUS BG245794 990 bp mRNA linear EST 13-FEB-2001
 DEFINITION 602358585F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4487133 5', mRNA sequence.
 ACCESSION BG245794
 VERSION BG245794.1 GI:12755609
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 990)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Arrayed by: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.

ORIGIN

Alignment Scores:
 Pred. No.: 3,766-26 Length: 990
 Score: 316.00 Matches: 53
 Percent Similarity: 81.82% Conservative: 10
 Best Local Similarity: 68.83% Mismatches: 14
 Query Match: 71.49% Indels: 0
 DB: 2 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BG245794 (1-990)

Qy 1 CysLysAspCysGlyLysAlaPheileGlnLysSerAsnLeuileArgHisGlnArgThr 20
 Db 114 TGTGACAGTGTGATAGGCAATTTGCACAAACAGTCATCTTAAACACATAAAGAAGACT 173
 Qy 21 HistHrGlyGluLysProTyrLysCysGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 174 CATACTGGAGAGAAACCCATGATGTGAGCAATGTGCAAAAGCCTTTGCAAGTAATAGT 233
 Qy 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 234 AATCTGCAAGTGCATAAAAAATGCACACTGGAGAGAAACCCATGATGTAAAGCAGTGT 293
 Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 294 GGTAAGCCTTTGGGTATCAAAAGTGGTCTGCAAAAGCATAAACGACACAT 344

RESULT 7

BG262051

LOCUS BG262051 1015 bp mRNA linear EST 13-FEB-2001
 DEFINITION 602373885F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481397 5', mRNA sequence.
 ACCESSION BG262051
 VERSION BG262051.1 GI:12771867
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1015)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1AM10316 row: j column: 22
High quality sequence start: 3
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

source
1. .1015
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4481397"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 3 88e-26 Length: 1015
Score: 316.00 Matches: 53
Percent Similarity: 81.82% Conservative: 10
Best Local Similarity: 68.83% Mismatches: 14
Query Match: 71.49% Indels: 0
DB: 2 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BG262051 (1-1015)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 174 TGTGACCATGTGATGAAGCATTTGGCAAAACAGTCATCTCTTAACATTAAGAACT 233
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 234 CATACTGGAGAGAAACCCCTATGAATGTGAGCAATGTGGCAAGCCCTTTCAGATTAATAGT 293
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 294 AATCTGCAAGTGCATAAAAAATGACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGT 353
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 354 GGTAAAGCCTTTGGGTATCAAGTGTCTGTGCAAAAGCATAAAGACACAT 404

RESULT 8
CF145979 536 bp mRNA linear EST 06-AUG-2003
LOCUS UI-HF-CB0-asj-f-01-0-UI.r1 NIH_MGC_210 Homo sapiens CDNA clone
DEFINITION IMAGE:30568668 5', mRNA sequence.
ACCESSION CF145979
VERSION CF145979.1 GI:33261423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 536)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548

COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa

375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers
1. .536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30568668"
/tissue_type="CNCAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC 210"
/note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pT7T3 Pac vector. The library tag sequence located between the Not I site and the polyA tail is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

Alignment Scores:
Pred. No.: 2 29e-26 Length: 536
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x CF145979 (1-536)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 90 TGTAAAGAATGTGGCAAGCATTTATGTCTTCAACCCCTAAGTACATAGAGAGATA 149
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 150 CACACTGGAGAGAACCCCTACAAATGTGAAGAAATGTGGCAAGCCTTTAGCGCTTCTCA 209
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 210 ACCCTTACTAGCATTAAGCAATTCATCTGGAGAGAAACCCCTACAAATGTAAGATGT 269
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 270 GGCAAGCTTTTAAAGCACTCTCTGAGCAATGTGTAACATAATAATACAT 320

RESULT 9

CF409262/c

LOCUS CB409262/2

DEFINITION NISC_nc04a08.x1 COGENE 6E MAX Homo sapiens CDNA clone IMAGE:5796015

3', mRNA sequence.

ACCESSION CB409262

VERSION CB409262.1 GI:29166131

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo;

REFERENCE 1 (bases 1 to 632)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation:
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM12900 row: B column: 16
Seq primer: -21M13 forward primer (ABI).

Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
 Isogai, T.
 HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Iehii, S.,
 Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T.,
 Suzuki, Y., Sugano, S., Isogai, T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers
 1..827
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RM200735"
 /cell_type="teratocarcinoma"
 /cell_line="NT2"
 /clone_lib="NT2RM2"
 /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
 precursor cells"

ORIGIN

Alignment Scores:	
Pred. No.:	2.81e-26
Score:	315.00
Percent Similarity:	81.82%
Best Local Similarity:	68.83%
Query Match:	71.27%
DB:	6
	Length: 632
	Matches: 53
	Conservative: 10
	Mismatches: 14
	Indels: 0
	Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x CB409262 (1-632)

Qy	1	CysLysaspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr	20
Db	280	TGTGAAGAATGTGGCAAAGCCTTTACATGTCTGCAGGCCTCCATAAACATAGAGAACT	221
Qy	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
Db	220	CATACTGGAGAGAAACCCCTACAAATGTGAAGAATGTGGCAAGCGTATACCTACCTCA	161
Qy	41	AsnLeuThrLysHisLysLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
Db	160	AATCTAACTGACACATAGACAACTCATCTACTGGAGAGAAACCTTACAANTGTAAGAATGT	101
Qy	61	GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis	77
Db	100	GGCAAGCCTTTTAACCTGGTCCTCCAGACCTTTAAATAAACATAGAGAAATTCAT	50

[illegible]

```

ORIGIN
precursor cells

Alignment Scores:
Pred. No.:      3_93e-26      Length:      827
Score:          315.00      Matches:      54
Percent Similarity: 84.42%      Conservative: 11
Best Local Similarity: 70.13%      Mismatches:  12
Query Match:    71.27%      Indels:      0
DB:             1      Gaps:          0

US-10-669-861-2_COPY_33_109 (1-77) x AU123664 (1-827)

Qy      1  CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
      Db      446  TGTGAAGAATGTGGCAAAAGCTTTTAAACCAAGTCCTCATCCCTTACTAAACATAAGAAAAATT 505

Qy      21  HisThrGlyGluLysProThrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      Db      506  CATACCTGGAGAGAAACCCCTACAAATGTGAAGAATGTGGCAAGCTTTTAAACCAAGTCCTCT 565

Qy      41  AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProThrLysCysLysGlnCys 60
      Db      566  TCCCTTACTTAAACATTAAGAAATCCATTCCTGGAGAGAAACCATACAGAGTGTGATAATGT 625

Qy      61  GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
      Db      626  GGCAAAAGCCTTTATTTTACCCTCAAGCCTTACTAGACATGAGATAATTTCAT 676

```

RESULT	11
BQ221286	
LOCUS	BQ221286 linear EST 02-MAY-2002
DEFINITION	AGENCOURT_7504692 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015192 5' mRNA sequence.
ACCESSION	BQ221286
VERSION	BQ221286.1 GI:20402686
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens <i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini, Hominoidea; Hominidae; Homo.</i>
REFERENCE	NH-GMC http://mgc.nci.nih.gov/ (bases 1 to 875) AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC) TITLE Unpublished (1999) JOURNAL

[illegible]

White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES Location/Qualifiers

source 1..1362
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"

gene <1..>1362
/locus_tag="HC1584"

ORIGIN

Alignment Scores:
Pred. No.: 7,3e-26 Length: 1362
Score: 315.00 Matches: 53
Percent Similarity: 81.82% Conservative: 10
Best Local Similarity: 68.83% Mismatches: 14
Query Match: 71.27% Indels: 0
DB: 11 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x DQ030160 (1-1362)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 1108 TGTGAAGATGTGGCAAGCCCTTTACATGGTCTGCAGGCCCTCCATAAACAATAGGAGAACT 1167
Qy 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 1168 CATACTGGAGAGAAACCTTACAAATGTGAAGATGTGGCAAGCCCTATACATCTCA 1227
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 1228 AATCTAACTGAACATAAGCAACTCATACTGGAGAGAAACCTTACAAATGTGAAGATGT 1287
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 1288 GGCAAGGCTTTTAACTGGTCTCAGACCTTAATAAACATAAGAGAAATTCAT 1338

RESULT 14
BX483171 441 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686B22235_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFZp686B22235 5', mRNA sequence.

ACCESSION BX483171
VERSION BX483171.1 GI:31943773

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 441)
Flockerzi,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and Wiemann,S.
EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS

MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

No si sequence available.
This clone (DKFZp686B22235) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

FEATURES source

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers
1..441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686B22235"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Alignment Scores:
Pred. No.: 2,35e-26 Length: 441
Score: 314.00 Matches: 54
Percent Similarity: 80.52% Conservative: 8
Best Local Similarity: 70.13% Mismatches: 15
Query Match: 71.04% Indels: 0
DB: 5 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BX483171 (1-441)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 124 TGCAGTGAATGTGTAAGCCCTTCATTCAGAGCATTTGCCCTTATTCGCACACAGAGAAGT 183
Qy 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 184 CACACTGGAGAGAAACCTATAAATGCAATGATGTGGAAAGGCTTTAATCAGAACACC 243
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 244 TGCCTCACTCATATGAGAAATTCATCTGGAGAGAGCCCTATAAATGTAAAGAATGT 303
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 304 GGGAAAGCCCTTGTCTCATGCTCATCTTACTGAACATCATAGAACTCAC 354

RESULT 15
DQ045809

LOCUS DQ045809 584 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes ZNF146 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION DQ045809
VERSION DQ045809.1 GI:66897024

KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.

REFERENCE 1 (bases 1 to 584)
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
(er) PLoS Biol. 3 (6), E170 (2005)

JOURNAL PLoS Biol. 3 (6), E170 (2005)

PUBMED 15869325

REFERENCE 2 (bases 1 to 584)

Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

Direct Submission

Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

TITLE This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES source

Location/Qualifiers
1..584

```
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>584
/gene="ZNF146"
/locus_tag="HC11576"
```

Gene

ORIGIN

Alignment Scores:

Pred. No.:	3.33e-26	Length:	584
Score:	314.00	Matches:	54
Percent Similarity:	80.52%	Conservative:	8
Best Local Similarity:	70.13%	Mismatches:	15
Query Match:	71.04%	Indels:	0
DB:	11	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x DQ045809 (1-584)

QY	1	CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuIleArgHisGlnArgThr	20
DB	304	TGTAAGATTGGCGAAGCTTTTCATTCAAGAGTCAACCTCATCAGACACGAGAACT	363
QY	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer	40
DB	364	CACACAGGAGAGAGCCCTTTGTATGTAAGAGTGTGMAAAACCTTCAGTGGCAATCC	423
QY	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
DB	424	AACCTTACTGAGCATGAGAAATCCATATTGGAGAGAGCCCTTTAAATGTAGTGAATGT	483
QY	61	GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis	77
DB	484	GGAACAGCCCTTGGCCAGAGAAGTACCTCATATAAACATCAAAACATTCCAC	534

Search completed: January 12, 2006, 03:29:25
 Job time : 3029 secs

This Page Blank (uspto)